

The Telecommunications and Data Acquisition Progress Report 42-92

October–December 1987

E. C. Posner
Editor

February 15, 1988

NASA

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

Preface

This quarterly publication provides archival reports on developments in programs managed by JPL's Office of Telecommunications and Data Acquisition (TDA). In space communications, radio navigation, radio science, and ground-based radio and radar astronomy, it reports on activities of the Deep Space Network (DSN) and its associated Ground Communications Facility (GCF) in planning, in supporting research and technology, in implementation, and in operations. Also included is TDA-funded activity at JPL on data and information systems and reimbursable DSN work performed for other space agencies through NASA. The preceding work is all performed for NASA's Office of Space Operations (OSO).

In geodynamics, the publication reports on the application of radio interferometry at microwave frequencies for geodynamic measurements. In the search for extraterrestrial intelligence (SETI), it reports on implementation and operations for searching the microwave spectrum. In solar system radar, it reports on the uses of the Goldstone solar system radar for scientific exploration of the planets, their rings and satellites, asteroids, and comets. These three programs are performed for NASA's Office of Space Science and Applications (OSSA).

Finally, tasks funded under the JPL Director's Discretionary Fund and the Caltech President's Fund which involve the TDA Office are included.

This and each succeeding issue of the TDA Progress Report will present material in some, but not necessarily all, of the following categories:

OSO Tasks

DSN Advanced Systems

- Tracking and Ground-Based Navigation
- Communications, Spacecraft-Ground
- Station Control and System Technology
- Network Data Processing and Productivity

DSN Systems Implementation

- Capabilities for Existing Projects
- Capabilities for New Projects
- New Initiatives
- Network Upgrade and Sustaining

DSN Operations

- Network Operations and Operations Support
- Mission Interface and Support
- TDA Program Management and Analysis

GCF Implementation and Operations

Data and Information Systems

OSSA Tasks:

Search for Extraterrestrial Intelligence

Geodynamics

- Geodetic Instrument Development

- Geodynamic Science

Goldstone Solar System Radar

Discretionary Funded Tasks

Contents

OSO TASKS DSN Advanced Systems TRACKING AND GROUND-BASED NAVIGATION

Source and Event Selection for Radio-Planetary Frame-Tie Measurements Using the Phobos Landers	151
R. Linfield and J. Ulvestad NASA Code 310-10-60-50-00	
The JPL Trapped Mercury Ion Frequency Standard	13 ⁵²
J. D. Prestage, G. J. Dick, and L. Maleki NASA Code 310-10-62-15-00	

COMMUNICATIONS, SPACECRAFT-GROUND

Using Ridge Regression in Systematic Pointing Error Corrections	20 ⁵³
C. N. Guiar NASA Code 310-20-65-63-00	
An Optimal Structure for a 34-Meter Millimeter-Wave Center-Fed BWG Antenna: The "Cross-Box" Concept	27 ⁵⁴
K. L. Chuang NASA Code 310-20-65-86-04	
A 2.3-GHz Maser at Usuda, Japan, for TDRSS-Orbiting VLBI Experiment	40 ⁵⁵
R. B. Quinn NASA Code 310-20-66-09-00	
A Lower Bound for the Decoder Error Probability of the Linear MDS Code	43 ⁵⁶
K. -M. Cheung NASA Code 310-30-71-83-02	
A Labeling Procedure for Linear Finite-State Codes	50 ⁵⁷
K. -M. Cheung NASA Code 310-30-71-83-02	
Further Results on Finite-State Codes	56 ⁵⁸
F. Pollara, K. -M. Cheung, and R. J. McEliece NASA Code 310-30-71-83-04	
A Comparison of VLSI Architectures for Time and Transform Domain Decoding of Reed-Solomon Codes	63 ⁵⁹
I. S. Hsu, T. K. Truong, L. J. Deutsch, E. H. Satorius, and I. S. Reed NASA Code 310-30-71-87-02	

STATION CONTROL AND SYSTEM TECHNOLOGY

Spacecraft Telecommunications System Mass Estimates	82 ⁵¹⁰
J. H. Yuen and L. L. Sakamoto NASA Code 310-30-71-83-04	

DSN Systems Implementation CAPABILITIES FOR NEW PROJECTS

Phase Calibration Generator	89 ⁵¹¹
E. H. Sigman NASA Code 314-40-51-32-09	

A Simplified, General-Purpose Deep-Space Ranging Correlator Design	105	<i>512</i>
R. C. Tausworthe and J. R. Smith NASA Code 314-40-32-10-11		
Frame Synchronization Performance and Analysis	110	<i>513</i>
C. S. R. Aguilera, L. Swanson, and G. H. Pitt III NASA Code 314-30-57-01-30		
A Noise-Adding Radiometer for the Parkes Antenna	117	<i>514</i>
T. J. Brunzie NASA Code 314-30-66-11-15		
X-Band System Performance of the Very Large Array	123	<i>515</i>
J. S. Ulvestad, G. M. Resch, and W. D. Brundage NASA Code 314-40-66-09-12		

NETWORK UPGRADE AND SUSTAINING

DSS 14 64-Meter Antenna S- and X-Band Efficiency and System Noise Temperature Calibrations, September 1987	138	<i>516</i>
S. D. Slobin NASA Code 314-40-56-44-14		
New Mode Switching Algorithm for the JPL 70-Meter Antenna Servo Controller	147	<i>517</i>
J. A. Nickerson NASA Code 314-30-56-44-19		
Optical Links in the Angle-Data Assembly of the 70-Meter Antennas	154	<i>518</i>
M. D. Nelson, J. R. Schroeder, and E. F. Tubbs NASA Code 314-30-41-70-21		
Seventy-Meter Antenna Performance Predictions: GTD Analysis Compared With Traditional Ray-Tracing Methods	166	<i>519</i>
J. M. Schredder NASA Code 314-40-41-82-09		
The Design and Implementation of the Technical Facilities Controller (TFC) for the Goldstone Deep Space Communications Complex	175	<i>520</i>
D. A. Killian, F. J. Menninger, T. Gorman, and P. Glenn NASA Code 314-40-44-14-23		

OSSA TASKS Search for Extraterrestrial Intelligence

SETI Data Controllers	184	<i>521</i>
R. M. Gosline NASA Code 199-50-62-10-02		
Simulated Performance of an Order Statistic Threshold Strategy for Detection of Narrowband Signals	191	<i>522</i>
E. Satorius, R. Brady, W. Deich, S. Gulkis, and E. Olsen NASA Code 199-50-67-09-05		
Author Index, 1987	198	

Source and Event Selection for Radio-Planetary Frame-Tie Measurements Using the Phobos Landers

R. Linfield and J. Ulvestad
Tracking Systems and Applications Section

The Soviet Phobos Lander mission will place two spacecraft on the Martian moon Phobos in 1989. Measurements of the range from Earth-based stations to the landers will allow an accurate determination of the ephemerides of Phobos and Mars. Delta VLBI between the landers and compact radio sources nearby on the sky will be used to obtain precise estimates of the angular offset between the radio and planetary reference frames. The accuracy of this frame-tie estimate is expected to be in the vicinity of 10 nrad, depending on how well several error sources can be controlled (calibrated or reduced). Many candidate radio sources for VLBI measurements have been identified, but additional work is necessary to select those sources which have characteristics appropriate to the present application. Strategies for performing the source selection are described below.

I. Introduction

The technique of Very Long Baseline Interferometry (VLBI) allows very precise navigation of spacecraft [1]. Radio transmissions from the spacecraft are received at two widely separated radio telescopes on the earth, and recorded on magnetic tape. The data from the two telescopes are later correlated. This allows the measurement of the delay in arrival time between the two telescopes of wavefronts from the spacecraft. This delay can be measured very accurately (much less than 1 nsec error). A compact extragalactic radio source nearby (typically 10 degrees or less away) on the sky is observed immediately before or after (or both before and after) the spacecraft observation. This procedure allows the location of the spacecraft on the sky to be measured very accurately (5-50 nanoradians) with respect to the radio source. Extensive VLBI observations at JPL over the last ten years have led

to the development of a catalog of compact radio sources over the northern 75 percent of the sky, with positions accurate to 5-15 nanoradians (nrad), as described in [2]. The set of radio sources in this catalog forms what is known as the "radio frame." Because of the enormous distances (mostly more than 1 gigaparsec: 3×10^{22} km) of the objects in this catalog (the majority of which are quasars), their angular motions are negligible. Therefore, the radio frame is believed to be inertial and is stable at the approximate level of the measurement accuracy.

A VLBI measurement of a spacecraft with respect to any one of these compact radio sources (Δ VLBI) establishes the spacecraft position in the radio frame to an accuracy δ , where δ is the vector sum of the spacecraft-radio source measurement error and the source position error in the radio source

catalog. The value of δ is as small as 15–25 nrad currently [1], with values of 5–10 nrad expected within 5 years. Most deep space navigation purposes require the location of a spacecraft with respect to a planet or natural satellite. To make optimum use of Δ VLBI for navigation, it is therefore necessary to measure the location of solar system objects in the radio frame, thus performing a “radio-planetary frame-tie.” Currently, the accuracy of this frame-tie is approximately 200 nrad for the inner planets, and somewhat poorer for the outer planets [3]. An improvement in this value is critical for high-accuracy target-relative navigation by Δ VLBI.

II. The Phobos Lander Mission

The Soviet Phobos Lander Mission, scheduled for dual launches in July 1988, will land two spacecraft on the Martian moon Phobos in the spring of 1989. The spacecraft lifetime is estimated at 2–3 years. Ranging measurements between Earth and the landers will determine the orbit of Phobos around Mars (orbital period 7 hr 39 min) very precisely (much better than 5 nrad as seen from the earth). In addition, these measurements will improve our knowledge of the Martian orbit, which has been most accurately measured by ranging to the two Viking landers. Ranging measurements to a spacecraft fixed on the surface of a planet or a natural satellite are more useful than ranging measurements to a free-flying spacecraft because non-gravitational forces on a planet or natural satellite are negligible. This greatly reduces the number of parameters and the magnitude of systematic errors in the orbit determination process. With ranging data to the Viking Landers starting in 1976 and similar data to the Phobos Landers as late as 1991 or 1992, secular and long-period terms in the orbit of Mars can be well measured. This should allow the Martian orbit to be determined to about 5 nrad relative to the orbit of the earth around the sun.

The orbits of the earth and Mars will then form a planetary frame which is as accurate and stable as the radio frame, except for a very slow rotation of the planetary frame due to unmodeled effects such as unknown asteroid masses [4]. The orbits of other planets will be tied to this frame to variable accuracy (15–30 nrad for Venus; 200–400 nrad for the gas giants).

In addition to ephemeris improvement, the Phobos Lander Mission also provides the opportunity for accurate radio-planetary frame-tie measurements. Each lander will have a “VLBI broadcast mode,” with two coherent tones spaced 14.71425 MHz apart on the downlink frequency of 1.7 GHz. Delta VLBI measurements of a lander and a compact radio source nearby on the sky can locate the lander and therefore Mars in the radio frame (this is possible because the positions of Phobos and Mars will be tied together to much less than 5 nrad). A cursory error analysis concludes that a measure-

ment with an accuracy of 5–10 nrad can be obtained with a lander-radio source separation of 2 degrees or less, if adequate accuracy can be achieved in the calibration of instrumental phase errors in the ground system.¹ The error analysis assumed one measurement each of the lander and radio source, closely spaced in time. This separation is much less than for most current Δ VLBI measurements. Phobos Lander transmissions will be at a single frequency, which precludes direct calibration of the ionosphere or solar plasma. Furthermore, that frequency (1.7 GHz) is sufficiently low that phase errors due to charged particles in the ionosphere or solar plasma are large. For Δ VLBI measurements these errors are reduced, approximately by the spacecraft-radio source separation in radians.

Restricting Δ VLBI measurements to lander-radio source separations of less than 2 degrees is the most obvious technique for achieving high accuracy. However, it is not the only possible technique. Larger separations, with well-studied compact radio sources, could be used if several error sources (e.g., the troposphere, or charged particles) can be accurately calibrated. One possible technique is the local network approach [5], in which observations are made of several compact radio sources that are near the spacecraft on the sky. This process allows several atmospheric, earth orientation, and clock parameters to be determined from the data. These techniques (other than that of using small lander-source separations) will not be discussed further in this article. However, a careful error analysis is needed to determine their potential accuracy for Phobos Lander frame-tie measurements. A hybrid technique with fairly small lander-source separations and some calibration methods may prove to be optimum.

If small lander-source separations are used, it is necessary to begin the source selection process now. The Soviets will fix the lander transmission times in November 1987. Close passes between Mars and radio sources last for approximately 1 day or less time, and we must therefore determine all potential sources soon. There will be time between November 1987 and the arrival at Phobos to decide which of these sources (and what strategies) to use. For sources far (>5 degrees) from the path of Phobos, the lander-source distance changes proportionately much more slowly. The scheduling of observations is therefore much more flexible.

III. Source Selection

The ideal compact radio source for VLBI astrometry would be very strong (flux density >10 Jy), with all its emission concentrated within a region on the sky much smaller than 1 nrad

¹C. E. Hildebrand, “First Cut at Phobos Lander VLBI Errors,” IOM 335.1-87-29 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 3, 1987.

diameter. In practice, there are no such sources. The extent to which actual radio sources deviate from this ideal introduces two types of error into VLBI measurements: statistical error and error due to source structure effects. The flux density (total power received at the earth per unit area and frequency interval) and the size of a radio source both influence the S/N of a VLBI measurement. For VLBI, the appropriate measure of source strength is the correlated flux density. Very approximately (usually within a factor of 2–3), the correlated flux density is the power received from a strip on the sky with a width one-half the fringe spacing and oriented perpendicular to the vector between the two telescopes (the baseline vector). For a pair of telescopes separated by a distance d , the fringe spacing is $\lambda/(d|\sin \theta|)$, where λ is the observing wavelength, and θ is the angle between the source direction and the baseline vector [6].

The ratio of correlated flux density to total flux density is defined as the visibility. The visibility is always between 0 and 1, with the larger values implying more compact sources. Many radio sources have visibilities of 0 on intercontinental baselines (source sizes much larger than the fringe spacing), a few have visibilities only slightly less than 1.0 (source sizes less than half a fringe spacing), and many others have values somewhere in between. The visibility is a function of baseline length, orientation, and observing frequency. It will not be the same on the two primary Deep Space Network (DSN) baselines (Spain–California and California–Australia). Actual visibilities can only be determined by VLBI observations, although other information (e.g., source spectral index or connected-element interferometry) can help in predicting visibilities.

The flux densities of compact radio sources vary with time, typically on time scales of a few years. At 8.4 GHz, this is an important effect, as sources can strengthen or weaken by a factor of 2 in 1–2 years. However, at 1.7 GHz, source variability is comparatively weak. In a period of 5 years, almost no sources will vary by more than 20 percent. Therefore, we need not worry about source variability for the Phobos Lander Mission.

The observing frequency for the Phobos Landers is 1.7 GHz. All of our VLBI catalog observations are at X-band (8.4 GHz) and S-band (2.3 GHz). Correlated flux densities of radio sources at 1.7 GHz will usually be within a factor of two of correlated flux densities at 2.3 GHz. For Phobos Lander frame-tie measurements, a source with 50 mJy correlated flux density provides an acceptably small statistical error, if we have two 70 m antennas and receivers with 56 MHz bandwidth. We have set a requirement of 100 mJy correlated flux density on intercontinental baselines, with a preference for sources with correlated flux densities of 200 mJy or greater.

In addition to reduction in S/N, a non-zero source size introduces systematic error. VLBI observations of a compact radio source will measure a source position that depends on baseline length and orientation. The dependence is complicated [7], but the magnitude of the effect can be estimated within a factor of 3–4 by making measurements of the correlated flux density at several hour angles on the baselines that will be used for frame-tie measurements. In general, sources with high visibilities (greater than about 0.5) have small source structure effects on astrometry: less than 5 nrad at 1.7 GHz. Sources with lower visibilities may or may not have significant structure problems; it depends on the details of the structure. Correlated flux density measurements can be done during astrometric experiments on these sources, although extrapolating the effect from 2.3 GHz to 1.7 GHz introduces some uncertainty. We expect to discard between 10 percent and 30 percent of the candidate sources that have adequate correlated flux density, because those sources have structures which could cause unacceptable systematic errors.

A potential problem related to the source structure effect discussed above is the dependence of source position on observing frequency. Physically, this is due to the fact that lower frequency radiation arises from regions of lower particle energy and lower magnetic fields. Because there are strong spatial gradients of these properties in compact radio sources, and because most compact radio sources are strongly asymmetrical, the emission centroid varies with emission frequency. This effect has been observed to be as large as 10–15 nrad for 3C 273 between 10.7 GHz and 1.7 GHz. This is an extreme case (perhaps the largest of any source in the sky), and for almost all sources the effect will be less than 2–3 nrad. A source like 3C 273 can be identified (and removed from the catalog) by its low visibility (≈ 0.1) on intercontinental baselines. We therefore expect that position shifts due to the lower observing frequency will not pose a major problem.

If the radio frame were based on astrometric observations at 1.7 GHz, this effect would not be important. However, the radio frame is based on 8.4 GHz observations (with 2.3 GHz observations for ionospheric calibration), so the position offset between 8.4 GHz and 1.7 GHz is important. More analysis is needed to estimate the magnitude of this effect, and to determine if it can be calibrated. It may be possible to hold the effect to an acceptable level by selecting very compact sources, which are identified by their high visibilities.

If a candidate source lies within 1–2 degrees of the path of Phobos on the sky and has acceptable correlated flux density and structure, it is suitable for a frame-tie measurement. However, before the final data analysis can be performed, it is necessary to determine the source position to 5–10 nrad by

astrometric VLBI observations. This is a time-consuming process, and requires a large amount of antenna time. With the advent of wide-bandwidth recording capability at the DSN sites, substantially less antenna time will be needed. We may be able to obtain 5–10 nrad source positions for all successful candidates in 1–2 years, given adequate manpower and antenna time. However, accurate astrometry of candidate sources remains a problem, and obtaining 10 nrad source positions for all of them will be very difficult. We will probably have to settle for less accurate positions, or concentrate on a subset of the candidate sources.

Sources in the DSN Astrometric Catalog [2] have adequate correlated flux density on intercontinental baselines (greater than 250 mJy at S-band), and have positions known to 5–15 nrad. Table 1 presents the sources from that catalog which lie within 2 degrees of the path of Phobos between March 15, 1989, and May 1, 1992. Because Phobos is always within 30 sec of Mars on the sky, the ephemeris of Mars was used for this and all other source searches presented here. In Table 1, the epoch and distance of closest approach are given, along with the source name. The angular separation between Mars and the Sun at this epoch is also given. The relevance of this quantity is discussed in the section "Selection Among Candidate Events." The angular rate of motion of Mars on the sky at the epoch of closest approach is also listed. This value allows the Mars–radio source separation at nearby times to be easily calculated by the Pythagorean rule. (The path of Mars on the sky can be approximated as a straight line for short times.)

For several technical reasons, as discussed in the section "Selection Among Candidate Events," many apparently suitable candidate events may not be usable for frame-tie measurements. It is therefore desirable to search additional radio source catalogs in order to get more candidate events.

As a result of searches for navigation sources for various JPL missions (Voyager, Galileo, and Magellan), an ecliptic catalog of compact radio sources has been developed [8], [9]. These sources are compact, with correlated flux densities greater than about 100 mJy at S-band on intercontinental baselines. They should nearly all have greater than 100 mJy correlated flux density at 1.7 GHz, and are suitable for Δ VLBI measurements. However, astrometric VLBI observations have not yet been performed on these sources, and their positions are known only to about 1 arc second. Table 2 presents sources from this catalog which lie within 1 degree of the path of Mars in the period March 15, 1989, to May 1, 1992. The correlated flux densities of these sources at S-band on intercontinental baselines are given. For sources with multiple observations, the range of observed correlated flux densities is given.

A third useful set of radio sources is the VLA Calibrator Catalog. These sources are strong (more than 0.5 Jy) and usually dominated by moderately compact components (less than 1 sec in size). They have positions known to 0.3 sec or better. Their correlated flux densities on intercontinental baselines will be less than the catalog flux density by an amount that cannot be predicted for any given source. However, the correlated flux density will be at least 30 percent of the catalog flux density in many cases. Therefore, many of these sources will be good candidates for Phobos Lander frame-tie measurements. Table 3 presents sources from this catalog (and from additional sets of sources which have been observed with the VLA), which lie within 1 degree of the path of Mars in the period March 15, 1989, to May 1, 1992. Their 1.7 GHz flux densities are listed.

A fourth useful catalog of radio sources is the MIT–Green Bank Catalog [10]. This consists of approximately 6000 sources in the declination band 0 to 20 degrees, with total flux density at 5 GHz greater than 60 mJy. For many of these sources, there is little or no information beyond the total source flux density and the source position (30 arc seconds 1σ uncertainty). Table 4 lists all sources from this catalog which lie within 30 arc minutes of the Martian trajectory from March 15, 1989, to May 1, 1992. Table 5 lists sources which lie between 30 arc minutes and one degree of the Martian trajectory for the same period. In both tables, only sources with a total 5 GHz flux density greater than 150 mJy are included, and the flux density is listed. Many of these sources will have compact components too weak to use for frame-tie measurements. Only interferometric observations can determine which are suitable.

Numerous sources are in more than one of the catalogs which were searched. Those sources (if they lie close to the orbit of Mars) are listed only once: with the first catalog in which they appear. (This is the catalog for which the most source information is known.)

IV. Selection Among Candidate Events

One parameter which is listed for all candidate events is the Mars–sun angle (the arc length on the sky between Mars and the sun) at the time of the event. Because the orbit of Mars lies outside the earth, and nearly in the plane of the ecliptic, this parameter uniquely specifies the solar impact parameter of the rays from both the radio source and the spacecraft. The solar impact parameter (smallest distance between the sun and the ray paths) determines to the first order the column density of solar plasma traversed by these rays. Due to the single frequency of the lander downlink, this plasma introduces measurement errors. These errors are estimated as 5 nrad average (with a large scatter) for a Mars–sun angle of 15 degrees,

and nearly always less than 5 nrad for Mars-sun angles greater than 40 degrees.¹ We will therefore discard candidate events with Mars-sun angles less than 15 degrees and will give greater weight in our data analysis to events with angles greater than 30 degrees. By avoiding all events with Mars-sun angles less than 45 or 60 degrees the effects of solar plasma would be negligible. However, this would eliminate measurements over a substantial fraction of the Martian orbit, and reduce our ability to distinguish systematic errors.

Although the ionosphere and solar plasma are expected to cause the greatest transmission media effects, the troposphere is also of concern.¹ For a 1 degree elevation difference, tropospheric errors are estimated at 5 nrad if the source elevation is 12 degrees or less at either antenna. The use of spacecraft-radio source separations of 30 arc minutes or less can greatly alleviate this problem. In addition, we should attempt to schedule observations when the spacecraft and radio source are at 15 degrees or higher elevation at both antennas. Calibrations with water-vapor radiometers may be able to reduce tropospheric errors by a factor of three.²

Subject to the restrictions of source structure effects (discussed above) and of transmission times (discussed below), all sources in Tables 1 and 2 with sun-Mars angles greater than 15 degrees are suitable for frame-tie observations. The sources in Tables 3-5 require additional screening. Intercontinental VLBI observations of these sources are needed to determine if their correlated flux densities at 1.7 GHz are greater than the 100 mJy requirement, and to determine the magnitude of source structure problems.

Although the antennas will have steerable beams, the Soviets want to minimize antenna movement. Therefore, the antennas will probably have fixed orientations for periods of days or weeks. Their beams will sweep across the earth once per orbital period of Phobos (7.65 hr). The gain will be variable during this sweep, with the time between 3 dB points being 45 minutes if the earth passes through the center of the beam. In addition, the landers will not always be broadcasting in VLBI mode, but will be used for ranging and telemetry much of the time. We do not yet have a schedule for transmissions from the landers, and may not have such a schedule until after the landings on Phobos. In combination with the limited mutual visibility of radio sources on intercontinental baselines, these lander broadcast restrictions will eliminate some candidate frame-tie events. It is important to select a surplus of candidate events, as a substantial fraction (perhaps 50 percent to 80 percent) will prove unusable.

One key quantity in any frame-tie measurement is the lander-radio source separation. As shown in IOM 335.1-87-29,¹ the static troposphere and the ionospheric variations can introduce errors greater than 5 nrad for elevation differences between lander and radio source as small as 1 degree. Because of the large longitude differences between the two stations of DSN intercontinental baselines, the elevation difference for at least one site will almost always be greater than one-half the spacecraft-radio source separation. Therefore, frame-tie measurements with spacecraft-radio source separations less than 1 degree are strongly preferred, and separations less than 30 arc minutes are desired.

The situation is more complicated than this, however. Unless the lander-radio source separation is less than the primary beam size of one antenna (10 min), the lander and radio source cannot be observed simultaneously. One will move across the sky due to sidereal motion (15 arc minutes per minute of time) while the other is being observed. Depending on the relative positions of the lander and radio source on the sky, this motion may increase or decrease the elevation difference. The observing sequence can be chosen to minimize the difference. Alternatively, a more complicated observing sequence (lander/radio source/lander or radio source/lander/radio source) can be used. This will allow a measurement of the linear component in the spatial and time variation in the troposphere and ionosphere.

In order to tie together two spherical coordinate systems, three parameters are needed. These can be thought of as resulting from the alignment of a reference point in the two systems (2 parameters), and then performing a rotation about this fixed point to align the remaining points (1 parameter). Two spacecraft-radio source measurements suffice, in principle, to determine these 3 parameters. However, if the two radio sources are close together (<20-30 degrees) or nearly opposite (>150-160 degrees) on the sky, the third parameter (rotation) is poorly determined. Two widely spaced measurements are needed. In practice, it is desirable to have many more than two measurements to reduce statistical measurement error. More importantly, multiple measurements over as much of the Martian trajectory as possible can allow an analysis of systematic errors, such as source structure effects, ionospheric perturbations, or even irregularities in either the planetary or radio reference frames.

V. Summary

Many candidate events for frame-tie measurements have been found. In order to determine which events to use, the following procedure will be adopted:

²S. E. Robinson, R. N. Treuhaft, B. L. Gary, and C. J. Vegos, "Tropospheric Wet Delay Calibrations for Magellan Navigation," IOM 335.3-87-79 (internal document), Jet Propulsion Laboratory, Pasadena, California, June 30, 1987.

- (1) Short-baseline VLBI observations will be performed on those sources not previously observed interferometrically. These observations will eliminate sources with sizes greater than 200 nrad, and will reduce the position errors of MIT-GB sources by a factor of about 100.
- (2) Sources that have survived this sieving process will be observed on intercontinental baselines at 1.7 or 2.3 GHz. This will determine the correlated flux density, and will allow an estimate of the source structure effects on astrometry.
- (3) Among sources close to the path of Mars, and which have sufficient correlated flux density, additional

screening will be done. Sources with low or variable visibilities (suggesting source structure problems) will be eliminated. Preference will be given to sources with the following properties: more than 200 mJy correlated flux density, a Mars-sun angle greater than 30 degrees, a Mars-radio source separation of less than 30 arc minutes, and a source position that is already well known from astrometric VLBI measurements. Additionally, an attempt will be made to select events which are uniformly spaced along the orbit of Mars.

- (4) Astrometric VLBI observations of the selected sources will be started. These observations will not be finished when some of the actual frame-tie observations are made, but will be needed for the final data analysis.

References

- [1] J. S. Border, F. F. Donovan, S. G. Finley, C. E. Hildebrand, B. Moultrie, and L. J. Skjerve, "Determining Spacecraft Angular Position with Delta VLBI: The Voyager Demonstration," AIAA Paper 82-1471, presented at the 1982 AIAA Conference, San Diego, California, August 1982.
- [2] J. L. Fanelow, O. J. Sovers, J. B. Thomas, G. H. Purcell, Jr., E. J. Cohen, D. H. Rogstad, L. J. Skjerve, and D. J. Spitzmesser, "Radio Interferometric Determination of Source Positions Utilizing Deep Space Network Antennas—1971 to 1980," *Astronomical Journal*, vol. 89, pp. 987-998, 1984.
- [3] A. E. Niell, XX Newhall, R. A. Preston, G. L. Berge, D. O. Muhleman, D. J. Rudy, J. K. Campbell, P. B. Esposito, and E. M. Standish, "Relating the Planetary Ephemerides and the Radio Reference Frame," *TDA Progress Report 42-81*, vol. January-March 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 1-8, May 15, 1985.
- [4] J. G. Williams, "Determining Asteroid Masses From Perturbations on Mars," *Icarus*, vol. 57, pp. 1-13, 1984.
- [5] R. N. Truhaft, "Astrometry in Local Reference Frames for Deep Space Navigation," in *Proc. of IAU Symposium No. 129*, 1987
- [6] E. B. Fomalont and M. C. H. Wright, "Interferometry and Aperture Synthesis," in *Galactic and Extra-Galactic Radio Astronomy*, New York and Amsterdam: Springer-Verlag, 1974, pp. 256-290
- [7] J. S. Ulvestad, "Effects of Source Structure on Astrometry and Geodesy," in *Proc. of IAU Symposium No. 129*, 1987
- [8] J. S. Ulvestad and R. P. Linfield, "The Search for Reference Sources for Δ VLBI Navigation of the Galileo Spacecraft," *TDA Progress Report 42-84*, vol. October-December 1985, pp. 152-163, Jet Propulsion Laboratory, Pasadena, California, February 15, 1986.

- [9] A. E. Wehrle, D. D. Morabito, and R. A. Preston, "Very Long Baseline Interferometry Observations of 257 Extragalactic Radio Sources in the Ecliptic Region," *Astronomical Journal*, vol. 89, pp. 336-341, 1984.
- [10] C. L. Bennett, C. R. Lawrence, B. F. Burke, J. N. Hewitt, and J. Mahoney, "The MIT-Green Bank 5 GHz Survey," *Astrophysical Journal Supplement*, vol. 61, pp. 1-104, 1986.

Table 1. Close passes (<2 degrees) between Mars and the DSN Astrometric Catalog, 3/15/89-5/1/92

Date	UT	Source	Closest Approach (arc min)	Sun-Planet Angle (deg)	Angular Speed of Mars (arc min/hr)
1989 Jun. 7	18:14	B2 0745+24	93	38	1.6
1989 Jul. 3	04:58	OJ 287	82	29	1.6
1989 Aug. 1	16:48	GC 1004+14	69	20	1.6
1989 Oct. 12	15:53	3C 279	30	4	1.6
1989 Nov. 4	16:03	P 1352-104	29	12	1.7
1989 Dec. 2	02:54	P 1504-167	21	21	1.7
1990 Jan. 26	21:03	1748-253	99	39	1.8
1990 Apr. 12	05:34	OX-173	34	59	1.9
1990 Apr. 15	23:47	OX-192	65	59	1.9
1990 Jun. 28	17:50	GC 0119+04	114	76	1.7
1990 Jul. 16	15:10	P 0201+113	68	81	1.7
1990 Aug. 26	06:31	0341+158	114	96	1.3
1991 May 17	02:38	B2 0745+24	77	59	1.4
1991 Jun. 13	01:54	OJ 287	74	49	1.5
1991 Jul. 13	15:32	GC 1004+14	68	38	1.5
1991 Sep. 24	09:43	3C 279	18	14	1.6
1991 Oct. 17	06:26	P 1352-104	43	7	1.7
1991 Nov. 13	08:28	P 1504-167	36	1	1.7
1992 Jan. 6	19:09	1748-253	87	18	1.8
1992 Mar. 19	17:27	OX-173	43	36	1.9
1992 Mar. 23	08:24	OX-192	76	37	1.9
1992 Apr. 23	04:47	P 2320-035	115	43	1.9

Table 2. Close passes (<1 degree) between Mars and the DSN Ecliptic Catalog, 3/15/89–5/1/92

Date	UT	Source	Closest Approach (arc min)	Sun-Planet Angle (deg)	Angular Speed of Mars (arc min/hr)	Correlated Flux Density at S-band (mJy)
1989 Mar. 20	02:02	0409+22	34	66	1.5	210
1989 Mar. 25	10:42	0423+233	23	64	1.5	170
1989 Apr. 28	18:41	0556+238	55	51	1.5	240
1989 Apr. 30	18:16	0601+244	19	51	1.5	100
1989 Jun. 14	20:39	GC 0802+21	30	35	1.6	240
1989 Jun. 29	05:35	GC 0839+18	54	30	1.6	420
1989 Sep. 19	14:15	P 1158+007	17	3	1.6	140
1989 Sep. 20	21:02	P 1203+011	34	3	1.6	110
1989 Sep. 26	18:43	1216-010	11	1	1.6	150
1989 Oct. 2	03:39	P 1229-021	12	1	1.6	120
1989 Nov. 5	12:58	P 1354-107	21	12	1.7	90
1989 Nov. 22	10:16	P 1437-153	20	18	1.7	110
1989 Nov. 24	20:52	P 1443-162	47	19	1.7	330
1989 Nov. 25	14:01	P 1445-16	28	19	1.7	(350,530)
1990 Apr. 6	20:52	P 2126-15	30	57	1.9	130
1990 Apr. 20	17:56	P 2208-137	41	61	1.9	(130,260)
1990 Apr. 26	04:06	2223-114	13	62	1.9	(250,400)
1990 May 6	09:41	2252-090	3	64	1.9	(140,380)
1990 Jun. 5	01:26	0013-00	5	71	1.8	(230,330)
1990 Jun. 17	06:53	P 0047+023	30	74	1.8	(140,210)
1990 Jul. 27	09:51	0229+13	34	84	1.6	(220,480)
1990 Aug. 15	19:16	CTA 21	9	91	1.4	122
1990 Sep. 21	20:11	P 0428+20	7	109	0.9	130
1990 Nov. 19	08:25	GC 0423+23	41	168	0.9	170
1990 Nov. 27	23:36	0409+22	26	178	0.9	210
1991 Feb. 8	10:22	0409+22	25	106	0.9	210
1991 Feb. 17	03:06	GC 0423+23	32	101	1.0	170
1991 Mar. 8	02:25	0459+252	21	90	1.2	240
1991 Apr. 4	22:39	0601+244	51	76	1.3	100
1991 May 24	17:40	GC 0802+21	44	56	1.5	240
1991 Sep. 1	07:19	P 1158+007	8	21	1.6	140
1991 Sep. 2	14:09	P 1203+011	43	21	1.6	110
1991 Sep. 8	12:28	1216-010	1	19	1.6	150
1991 Sep. 10	09:11	P 1218-02	59	19	1.6	120
1991 Sep. 13	21:35	P 1229-021	22	18	1.6	120
1991 Oct. 18	03:09	P 1354-107	34	7	1.7	90
1991 Nov. 1	18:42	P 1430-155	52	2	1.7	(120,230)
1991 Nov. 3	19:32	P 1437-153	5	1	1.7	110
1991 Nov. 6	05:16	P 1443-162	33	1	1.7	330
1991 Nov. 6	22:08	P 1445-16	14	1	1.7	(350,530)
1992 Mar. 14	13:29	P 2126-15	23	35	1.9	130
1992 Mar. 27	22:41	P 2208-137	55	38	1.9	(130,260)
1992 Apr. 2	04:30	2223-114	3	39	1.9	(250,400)
1992 Apr. 12	01:01	2252-090	24	41	1.9	(140,380)

Table 3. Close passes (<1 degree) between Mars and VLA Catalog sources, 3/15/89-5/1/92

Date	UT	Source	Closest Approach (arc min)	Sun-Planet Angle (deg)	Angular Speed of Mars (arc min/hr)	Sub-arc-sec Component Flux Density (mJy)
1989 Jul. 6	02:36	MG0900+1831	15	28	1.6	110*
1989 Aug. 28	04:34	MG1109+0659	31	11	1.6	200*
1989 Sep. 15	04:57	MG1150+0115	40	5	1.6	130*
1989 Sep. 21	20:01	MG1208+0054	51	3	1.6	200*
1989 Oct. 7	05:16	1242-047	56	3	1.6	3400‡
1989 Nov. 24	20:45	1445-164	47	19	1.7	510†
1990 Jan. 5	12:15	1646-224	2	32	1.8	1900‡
1990 Jun. 7	07:17	0022+002	4	71	1.8	2700†
1990 Jul. 18	20:36	MG0211+1051	9	82	1.7	200*
1991 Jun. 16	03:03	MG0900+1831	8	48	1.5	110*
1991 Aug. 9	16:20	MG1109+0659	35	29	1.6	200*
1991 Aug. 27	21:27	MG1150+0115	32	23	1.6	130*
1991 Sep. 18	23:22	1242-047	45	16	1.6	3400‡
1991 Nov. 1	18:31	1433-158	52	2	1.7	400†
1991 Dec. 17	00:19	1646-224	12	12	1.8	1900‡

*Flux density in sub-arc-second component at 5 GHz.

†Flux density in sub-arc-second component at 1.7 GHz.

‡Flux density at 1.7 GHz, but not all in a compact component.

Table 4. Close passes (<30 arc min) between Mars and MIT-GB Catalog, 3/15/89-5/1/92

Date	UT	Source	Closest Approach (arc min)	Sun-Planet Angle (deg)	Total Flux Density at 5 GHz (mJy)	Angular Speed of Mars (arc min/hr)
1989 Jul. 9	17:33	MG0909+1751*	15	27	164	1.6
1989 Jul. 11	12:56	MG0914+1715*	1	26	458	1.6
1989 Jul. 12	16:55	MG0917+1717*	15	26	159	1.6
1989 Jul. 24	08:50	MG0945+1428*	14	22	237	1.6
1989 Jul. 26	03:57	MG0950+1419*	1	22	1451	1.6
1989 Jul. 26	16:38	MG0950+1344	29	21	163	1.6
1989 Aug. 2	05:48	MG1007+1248*	4	19	417	1.6
1989 Aug. 3	14:15	MG1010+1235*	9	19	164	1.6
1989 Aug. 5	10:30	MG1015+1227	28	18	292	1.6
1989 Aug. 14	11:47	MG1036+0956*	2	15	470	1.6
1989 Aug. 25	21:51	MG1104+0730	27	12	231	1.6
1989 Sep. 3	17:48	MG1124+0456*	5	9	393	1.6
1989 Sep. 7	20:38	MG1134+0358*	12	7	151	1.6
1989 Sep. 11	10:10	MG1142+0235*	17	6	172	1.6
1989 Sep. 24	10:22	MG1213-0013*	21	2	225	1.6
1990 Jun. 11	15:56	MG0034+0118*	13	72	208	1.8
1990 Jun. 21	12:08	MG0100+0417*	4	74	163	1.8
1990 Jun. 30	21:56	MG0125+0617*	26	77	206	1.7
1990 Jul. 3	04:30	MG0131+0703*	15	78	206	1.7
1990 Jul. 5	04:13	MG0135+0810*	26	78	704	1.7
1990 Jul. 10	09:04	MG0148+0927*	26	79	150	1.7
1990 Jul. 21	00:36	MG0217+1103*	27	82	480	1.6
1990 Jul. 22	11:42	MG0220+1121*	29	83	481	1.6
1990 Jul. 28	17:37	MG0235+1304	2	85	225	1.6
1990 Aug. 1	03:21	MG0244+1320	27	86	259	1.6
1990 Aug. 26	04:41	MG0342+1736*	14	95	162	1.3
1990 Aug. 28	17:40	MG0347+1749*	20	97	285	1.3
1990 Sep. 2	06:17	MG0356+1900*	19	99	371	1.2
1990 Sep. 5	04:49	MG0402+1929*	29	100	348	1.2
1991 Jun. 19	21:46	MG0909+1751*	9	46	164	1.5
1991 Jun. 21	19:00	MG0914+1715*	7	46	458	1.5
1991 Jun. 23	00:02	MG0917+1717*	10	45	159	1.5
1991 Jul. 5	02:02	MG0945+1428*	17	41	237	1.5
1991 Jul. 6	22:26	MG0950+1419*	1	40	1451	1.5
1991 Jul. 14	05:08	MG1007+1248*	4	38	417	1.5
1991 Jul. 15	14:25	MG1010+1235*	9	37	164	1.5
1991 Jul. 17	11:43	MG1015+1227	28	37	292	1.5
1991 Jul. 26	18:01	MG1036+0956*	4	34	470	1.6
1991 Aug. 16	07:39	MG1124+0456*	11	27	393	1.6
1991 Aug. 20	11:31	MG1134+0358*	19	26	151	1.6
1991 Aug. 24	01:54	MG1142+0235*	10	24	172	1.6

*VLA observations made after this article was submitted have shown that these sources have inadequate compact flux density for Δ VLBI observations.

Table 5. Close passes (30–59 arc min) between Mars and MIT–GB Catalog, 3/15/89–5/1/92

Date	UT	Source	Closest Approach (arc min)	Sun–Planet Angle (deg)	Total Flux Density at 5 GHz (mJy)	Angular Speed of Mars (arc min/hr)
1989 Jul. 5	01:48	MG0856+1739*	50	28	150	1.6
1989 Jul. 9	11:38	MG0909+1821*	44	27	306	1.6
1989 Jul. 10	13:42	MG0910+1650*	39	27	158	1.6
1989 Jul. 15	22:30	MG0925+1658	35	25	191	1.6
1989 Jul. 30	01:18	MG1000+1401*	36	20	514	1.6
1989 Jul. 31	22:09	MG1002+1215	50	20	284	1.6
1989 Aug. 1	12:19	MG1004+1207	51	20	167	1.6
1989 Aug. 8	02:01	MG1020+1039*	48	17	166	1.6
1989 Aug. 12	18:30	MG1034+1112*	58	16	464	1.6
1989 Aug. 13	17:43	MG1036+1052	52	16	201	1.6
1989 Aug. 19	16:33	MG1050+0926*	52	14	153	1.6
1989 Aug. 28	22:00	MG1109+0543	39	11	184	1.6
1989 Sep. 6	08:15	MG1131+0456*	50	8	205	1.6
1989 Sep. 12	03:42	MG1142+0154*	50	6	218	1.6
1989 Sep. 21	03:09	MG1204–0029*	53	3	184	1.6
1990 Jun. 18	13:08	MG0051+0358*	36	74	234	1.8
1990 Jun. 23	04:42	MG0103+0521	43	75	183	1.8
1990 Jul. 2	22:22	MG0131+0623*	53	77	545	1.7
1990 Jul. 17	06:38	MG0205+1134*	59	81	186	1.7
1990 Jul. 20	04:47	MG0213+1212	57	82	180	1.6
1990 Jul. 24	07:57	MG0225+1134	40	83	387	1.6
1990 Jul. 30	19:12	MG0241+1253*	39	86	189	1.6
1990 Jul. 30	20:40	MG0242+1248	45	86	161	1.6
1990 Aug. 11	06:24	MG0307+1609*	36	89	170	1.5
1990 Aug. 13	12:09	MG0314+1508*	51	90	187	1.5
1990 Aug. 23	20:51	MG0338+1634*	58	94	166	1.3
1990 Aug. 31	00:16	MG0352+1754	32	98	248	1.3
1990 Sep. 10	06:15	MG0412+1856	36	103	248	1.1
1991 Jun. 15	01:27	MG0856+1739*	58	48	150	1.5
1991 Jun. 19	15:27	MG0909+1821*	38	46	306	1.5
1991 Jun. 20	19:00	MG0910+1650*	45	46	158	1.5
1991 Jun. 23	18:22	MG0920+1753	56	45	353	1.5
1991 Jun. 26	08:31	MG0925+1658	30	44	191	1.5
1991 Jul. 7	11:38	MG0950+1344	31	40	163	1.5
1991 Jul. 10	22:24	MG1000+1401*	35	39	514	1.5
1991 Jul. 12	20:50	MG1002+1215	51	38	284	1.5
1991 Jul. 13	11:24	MG1004+1207	52	38	167	1.5
1991 Jul. 20	05:03	MG1020+1039*	47	36	166	1.5
1991 Jul. 25	23:24	MG1036+1052	54	34	201	1.5
1991 Aug. 1	01:00	MG1050+0926*	55	32	153	1.6
1991 Aug. 7	08:51	MG1104+0730	31	30	231	1.6
1991 Aug. 9	15:43	MG1107+0533*	59	29	285	1.6
1991 Aug. 10	10:13	MG1109+0543	35	29	184	1.6
1991 Aug. 15	00:23	MG1119+0410	59	27	235	1.6
1991 Aug. 18	22:39	MG1131+0456*	56	26	205	1.6
1991 Aug. 24	19:40	MG1142+0154*	42	24	218	1.6
1991 Sep. 1	01:29	MG1159–0015*	59	22	225	1.6
1991 Sep. 2	20:30	MG1204–0029*	44	21	184	1.6
1991 Sep. 6	03:53	MG1213–0013*	31	20	225	1.6

*VLA observations made after this article was submitted have shown that these sources have inadequate compact flux density for Δ VLBI observations.

The JPL Trapped Mercury Ion Frequency Standard

J. D. Prestage, G. J. Dick, and L. Maleki
Communications Systems Research Section

In order to provide frequency standards for the DSN which are more stable than present-day hydrogen masers, a research task was established under the Advanced Systems Program of the TDA to develop a $^{199}\text{Hg}^+$ trapped ion frequency standard. This article describes the first closed-loop operation of this standard.

Mercury 199 ions are confined in an RF trap and are state-selected through the use of optical pumping with 194-nm UV light from a ^{202}Hg discharge lamp. Absorption of microwave radiation at the hyperfine frequency (40.5 GHz) is signaled by atomic fluorescence of the UV light. The frequency of a 40.5-GHz oscillator is locked to a 1.6-Hz-wide atomic absorption line of the trapped ions. The measured Allan variance of this locked oscillator is currently $\sigma_y(\tau) = 4.4 \times 10^{-12}/\sqrt{\tau}$ for $20 < \tau < 320$ seconds, which is better stability than the best commercial cesium standards by almost a factor of 2. This initial result was achieved without magnetic shielding and without regulation of ion number.

I. Introduction

Recently, there has been much activity directed toward the development of trapped ion frequency standards. This has occurred because ions confined in an RF quadrupole trap are subjected to very small perturbations of their atomic energy levels and to weak forces which equalize any population differences among the ground-state hyperfine levels. Potentially, the largest source of frequency fluctuation for such a standard stems from the motion of the atoms within the trap via the second-order doppler or relativistic time dilation effect. To minimize this perturbation, heavy ions are preferable to light ions, since for a given energy a heavy ion will have a smaller velocity. For this reason and for other reasons discussed later in this article, $^{199}\text{Hg}^+$ ions have been used for most trapped ion frequency standard work.

Only two parameters are needed to describe the short-term stability of a passive atomic frequency standard. One is the

line Q, which equals $f/\Delta f$ where f is the resonant frequency of the reference atom and Δf is the width of the atomic resonance. For $^{199}\text{Hg}^+$, f is 40.5 GHz and Δf is as small as 0.1 Hz; thus, trapped ion standards have line Q's which are orders of magnitude higher than other microwave atomic frequency standards.

The other parameter that determines stability is the signal-to-noise ratio (SNR) achieved in measuring the atomic resonance. The short-term stability is inversely proportional to the product of Q and SNR. At present, four groups are developing trapped Hg^+ ion clocks: Hewlett-Packard in Palo Alto, California [1]; the National Bureau of Standards in Boulder, Colorado [2]; l'Universite Paris-Sud in Orsay, France [3]; and JPL/NASA. It should be noted that F. G. Major first proposed the use of trapped $^{199}\text{Hg}^+$ ions as a frequency standard while working for NASA at the Goddard Space Flight Center in 1969 [4].

II. Ion Trapping

The electrode structure and time-varying voltages used to trap ions are shown in Fig. 1. The electric potential inside the trap when no ions are present is

$$V_T = \frac{(U_0 + V_0 \cos \Omega t)(r^2 - 2z^2)}{\xi^2}$$

where, for the present work, $\Omega = (2\pi) 500$ kHz, $U_0 = 20$ V, $V_0 = 660$ V, $\xi^2 = (r_0^2 + 2z_0^2)$, and r_0 = the inside radius of ring electrode = 1.9 cm = $\sqrt{2} z_0$.

A charged particle moving in the inhomogeneous oscillatory electric field of the ion trap feels a net force (averaged over one cycle of Ω) toward the region of the weaker field provided that the amplitude of its motion at frequency Ω is small compared with its distance from the center of the trap. The motion under these conditions is a combination of a fast oscillation at frequency Ω (micromotion) and a slower frequency ω , as shown in Fig. 2.

The action of the RF field in trapping ions is described by the electric pseudopotential energy [5]:

$$\Psi = \frac{e^2 V_0^2 (r^2 + 4z^2)}{m\Omega^2 \xi^4}$$

where m and e are the ion's mass and charge, respectively. The dc potential energy

$$\phi_{dc} = \frac{eU_0 (r^2 - 2z^2)}{\xi^2}$$

is added to the pseudopotential, giving the total potential energy for an ion in the trap:

$$\phi_T = \frac{m\omega_r^2 r^2}{2} + \frac{m\omega_z^2 z^2}{2}$$

where

$$\omega_r^2 = \frac{2e^2 V_0^2}{m^2 \Omega^2 \xi^4} + \frac{2eU_0}{m\xi^2}$$

and

$$\omega_z^2 = \frac{8e^2 V_0^2}{m^2 \Omega^2 \xi^4} - \frac{4eU_0}{m\xi^2}$$

Under the conditions listed earlier, ≈ 19 electron volts of kinetic energy is required for a $^{199}\text{Hg}^+$ ion at the trap center to reach one of the trapping electrodes.

Our trap is inside a vacuum chamber with a pressure of $\approx 1-2 \times 10^{-8}$ torr. By heating a powder of isotopically enriched mercuric oxide (HgO) to about 100°C , a vapor of neutral ^{199}Hg fills the vacuum chamber to about 10^{-8} torr partial pressure. Electrons from an LaB_6 single-crystal filament are injected into the trap ($\approx 30 \mu\text{A}$, 300 V), ionizing some of the neutral ^{199}Hg inside the trap electrode structure.

The resulting ion cloud is much hotter than the room-temperature neutral vapor in part because ionization of the vapor takes place throughout the trap. It has been found experimentally that the average kinetic energy of the ions is about 10 percent of the well depth. The resulting 2 eV of kinetic energy would produce a fractional second-order doppler shift of about 10^{-11} . To reduce this shift, the vacuum system is filled to about 10^{-6} torr of ^4He . The $^{199}\text{Hg}^+$ ions collide with these room-temperature helium atoms and are cooled to just above room temperature.

III. Magnetic Levels and State Selection

The magnetic structure of the ground-state hyperfine levels of $^{199}\text{Hg}^+$ is shown in Fig. 3. The energy difference between the ($F = 0, m_F = 0$) and ($F = 1, m_F = 0$) levels is used to define the standard frequency, approximately $f_{\text{Hg}^+} = 40.507347997$ GHz. The measured frequency, f , depends quadratically on the magnetic field at the position of the ion cloud, $f = f_{\text{Hg}^+} + 97B^2$ (Hz/G²). For comparison, the field dependence for hydrogen atoms is $f = f_H + 2750B^2$ (Hz/G²).

The ions are state selected by use of optical pumping with light from a ^{202}Hg discharge lamp. The energy levels of $^{202}\text{Hg}^+$ and $^{199}\text{Hg}^+$ are compared in Fig. 4. Ultraviolet light of wavelength 194.2 nm (≈ 6.4 eV) from the ^{202}Hg lamp, when collected and focused onto the $^{199}\text{Hg}^+$ ions, will excite the transition $^2S_{1/2} (F = 1, m_F) \rightarrow ^2P_{1/2}$.

The $^2P_{1/2}$ state decays after 2 ns lifetime to either the $^2S_{1/2} (F = 0, m_F = 0)$ or the $^2S_{1/2} (F = 1, m_F)$ state, thereby scattering a 194-nm photon. Since the transition $^2S_{1/2} (F = 0, m_F = 0) \rightarrow ^2P_{1/2}$ is not resonant with the light from the ^{202}Hg lamp, the ions are pumped out of the $^2S_{1/2} (F = 1, m_F)$ states into the $^2S_{1/2} (F = 0, m_F = 0)$ state, at which time

the ions stop scattering UV light. A flux of about 3×10^{12} photons per second per square centimeter passing through the ion cloud will depopulate the $^2S_{1/2}$ ($F = 1, m_F$) levels in about 1/2 second. An oscillating magnetic field (strength 10^{-6} G) at a frequency of 40.507347997 GHz will transfer the atoms from the $^2S_{1/2}$ ($F = 0, m_F = 0$) state to the $^2S_{1/2}$ ($F = 1, m_F = 0$) state in about 1 second. The ions will then scatter UV light until they are pumped back into the nonfluorescing $^2S_{1/2}$ ($F = 0, m_F = 0$) state.

IV. UV Optical System

Figure 5 shows the optical system used to collect and focus UV light from the ^{202}Hg lamp onto the ion cloud. The lamp is excited with 15–20 watts of RF power (160 MHz), creating a very bright discharge in the quartz cell containing the ^{202}Hg vapor with about 30 millitorr of argon buffer gas. The useful light from the ^{202}Hg lamp is from the 194.2-nm transition in $^{202}\text{Hg}^+$, as shown in Fig. 4.

However, the brightest wavelengths produced in the lamp are from transitions in the neutral mercury atom. Any light detected at wavelengths other than 194 nm will degrade the SNR of the measured atomic resonance. There are three ways we suppress the detection of light with wavelengths different from 194 nm:

- (1) The photomultiplier tube (PMT) is sensitive only to light that has wavelengths between 160 and 320 nm with peak sensitivities of 12 to 15 percent at 210 nm. The brightest line coming from the lamp in this bandwidth is 254 nm and is 200 times brighter than the 194-nm line.
- (2) The ellipsoidal collection mirror has a thin-film dielectric coating which maximizes reflection at 194 nm while keeping the reflectivity at 254 nm at 10 percent, with lower reflectivities for longer wavelengths. The entrance window to the trap region is coated to be 98 percent reflective at 254 nm while being 90 percent transmitting at 194 nm.
- (3) All light collected by the detection optics—stray scattered light plus fluorescent light from the atoms—is filtered with a 194-nm bandpass filter with a peak transmission of 30 to 40 percent and a bandwidth of 45 nm. The solid angle subtended by the collection optics around the trap center is 5 percent of the total 4π solid angle. The total collection efficiency of this detection system is equal to the solid angle multiplied by the bandpass filter loss multiplied by the PMT sensitivity, or $0.05 \times 0.35 \times 0.12 = 2 \times 10^{-3}$.

The atomic fluorescence and stray scattered light can be seen in Fig. 6. The scan is triggered by the start of the electron pulse, which forms the ions and causes some increase in detected light. After about 1 second the electron pulse is switched off, the ions are pumped into the $^2S_{1/2}$ ($F = 0, m_F = 0$) level, and only stray light is collected. At 2 seconds, the 40.5-GHz radiation is switched on, transferring some of the atoms into the $^2S_{1/2}$ ($F = 1, m_F = 0$) state, where they scatter light as discussed earlier. Finally, at about 3 seconds, the microwaves are switched off and the atomic fluorescence dies away with a time constant of about 1/5 second.

V. Closed-Loop Operation

The sequence of operations used to carry out a measurement of the $^2S_{1/2}$ ($F = 0, m_F = 0$) \rightarrow $^2S_{1/2}$ ($F = 1, m_F = 0$) frequency is shown in Fig. 7. By repeating this sequence as the frequency of the microwaves is stepped in 0.2-Hz increments, we measure the resonance curve shown in Fig. 8. For the 0.5-second square microwave pulse used in this measurement, the smallest linewidth theoretically attainable is 1.6 Hz, which corresponds to $Q = 2.5 \times 10^{10}$.

Locking the 40.5-GHz oscillator to the 1.6-Hz-wide resonance line is done by stepping the oscillator 0.8 Hz to either side of the resonance and adjusting the center frequency of this ± 0.8 -Hz step to null the difference in fluorescence rates. More precisely, suppose the center frequency of the 40.5-GHz oscillator, F_i , is within one linewidth of the resonance. Three measurements of the fluorescence are made on alternate sides of F_i : C1 at $F_i + 0.8$ Hz, C2 at $F_i - 0.8$ Hz, and C3 at $F_i + 0.8$ Hz. The oscillator center frequency is then changed to

$$F_{i+1} = F_i + \frac{(0.8 \text{ Hz}/T)(C1 + C3 - 2C2)}{2(\text{SIGNAL})}$$

where SIGNAL is the height of the fluorescence above background and T is the loop time constant in units of measurement cycle time. The “second difference” ($C1 + C3 - 2C2$) is used because it is insensitive to linear drifts in lamp intensity. If the first difference ($C1 - C2$) had been used to change F_i , a linear drift in lamp intensity would have forced a linear drift in the 40.5-GHz oscillator’s frequency away from the atomic line center. The sequence of center frequencies obtained in this closed-loop operation is shown in Fig. 9. The measurement shown lasted just over 2 hours. A single measurement cycle—consisting of an electron pulse to load ions in the trap, a waiting period while ions are optically pumped, lamp off while microwaves drive the transition, and lamp on and counter on to monitor fluorescence—lasts about 2.5 seconds. The loop time constant, T , is 5 measurement cycles, i.e., about 12.5 seconds.

The Allan variance derived from the sequence of frequencies $\{F_i\}$ is shown in Fig. 10. The reference frequency for the 40.5-GHz oscillator is provided by a hydrogen maser (SAO-21 in the Frequency Standards Laboratory). For times that are long compared to the loop time constant, the Allan variance falls as $4.4 \times 10^{-12}/\sqrt{\tau}$. This short-term stability is nearly a factor of 2 better than that of the best commercial cesium standards ($8.5 \times 10^{-12}/\sqrt{\tau}$). These initial results were obtained without shielding the 0.8 G ambient magnetic field in the trap. The residual field dependence at 0.8 G is 157 Hz/G. A frequency stability of 2×10^{-13} at this field sensitivity requires magnetic field fluctuations smaller than 60 μ G over the 320 seconds required to reduce statistical error to 2×10^{-13} . For this reason, we have not pushed the Allan variance data beyond 320 seconds in this first test.

VI. Conclusion and Summary

In its first closed-loop operation, the frequency stability of the trapped $^{199}\text{Hg}^+$ frequency standard has been measured to be

$$\sigma_y(\tau) = \frac{4.4 \times 10^{-12}}{\sqrt{\tau}}$$

for $20 < \tau < 320$ seconds.

Many improvements are under way to increase the short- and long-term stability of this standard, including increased fluorescence collection efficiency, shielding the ambient magnetic field, and designs for traps which could store one hundred times the present ion number.

Acknowledgments

We wish to thank T. Tucker for mercury lamp construction and for assisting in the development of the 40.5-GHz oscillator.

References

- [1] L. S. Cutler, R. P. Giffard, P. J. Wheeler, and G. M. Winkler, "Initial Operational Experience with a Mercury Ion Storage Frequency Standard," in *Proceedings of the 41st Annual Symposium on Frequency Control*, pp. 12-19, May 1987.
- [2] D. J. Wineland, W. M. Itano, J. C. Bergquist, and F. L. Walls, "Proposed Stored $^{201}\text{Hg}^+$ Ion Frequency Standards," in *Proceedings of the 35th Annual Symposium on Frequency Control*, pp. 602-611, May 1981.
- [3] M. Jardino, M. Desaintfusien, and F. Plumelle, "Prospects for a Mercury Ion Frequency Standard," *Journal de Physique*, vol. C8, p. 327, 1981.
- [4] F. G. Major, "Remarks on the Application of RF Spectroscopy of Stored Ions to Frequency Standards," in *Proceedings of the FAA Conference on Utilization of Time and Frequency in Collision Avoidance Systems*, pp. 125-126, 1968.
- [5] H. G. Dehmelt, "Radio Frequency Spectroscopy of Stored Ions," *Adv. At. Mol. Phys.*, vol. 3, pp. 153-154, 1967.

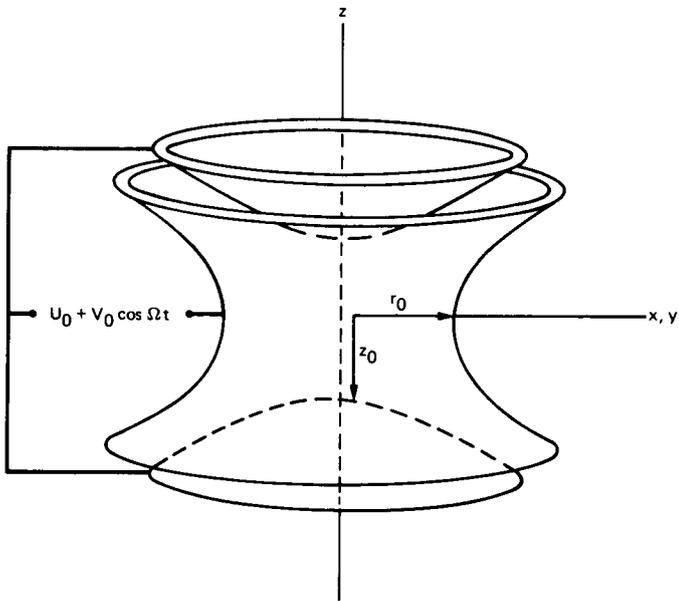


Fig. 1. Electrode structure and voltages used in this work

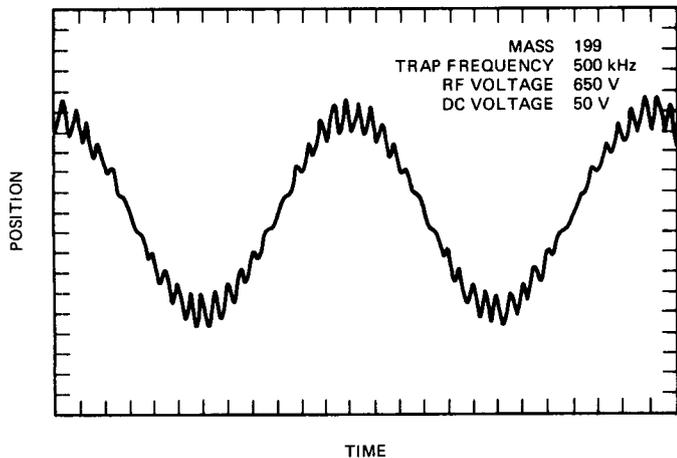


Fig. 2. One-dimensional motion for ion in combined RF and dc trapping fields

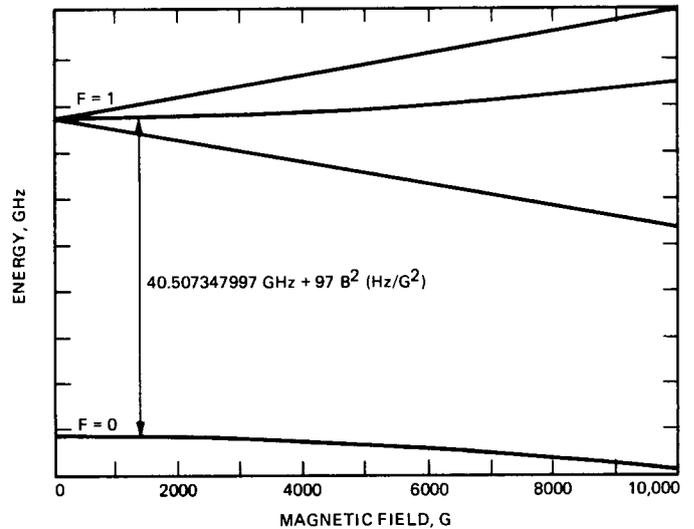


Fig. 3. Energy levels of the ground state of $^{199}\text{Hg}^+$ versus the magnetic field, showing the first-order field-independent transition used in this work

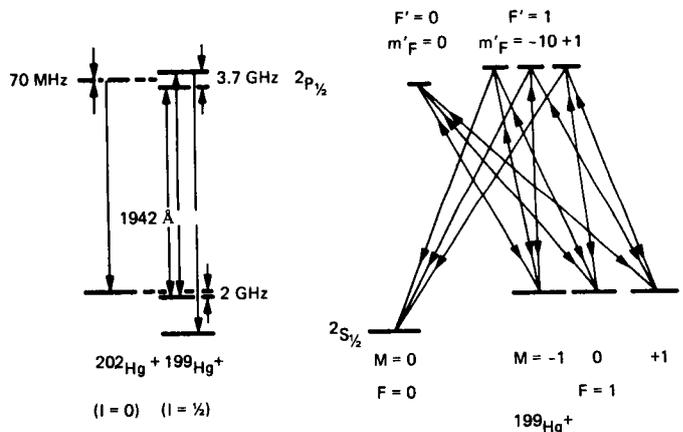


Fig. 4. Ground and lowest optically excited states of $^{202}\text{Hg}^+$ and $^{199}\text{Hg}^+$ (left); a depiction of the manner in which light from $^{202}\text{Hg}^+$ pumps $^{199}\text{Hg}^+$ ions into the $2S_{1/2}$ ($F=0, m_F=0$) level (right)

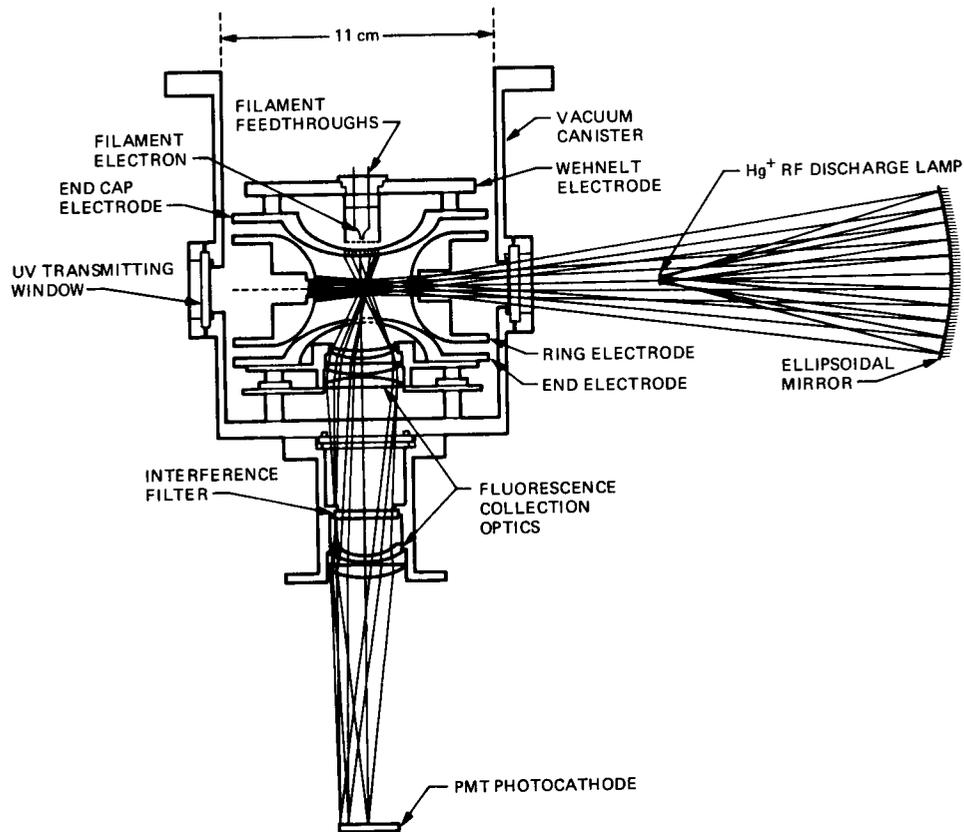


Fig. 5. UV light of 194 nm is scattered by the trapped $^{199}\text{Hg}^+$ ions after the absorption of microwave radiation near 40.5 GHz

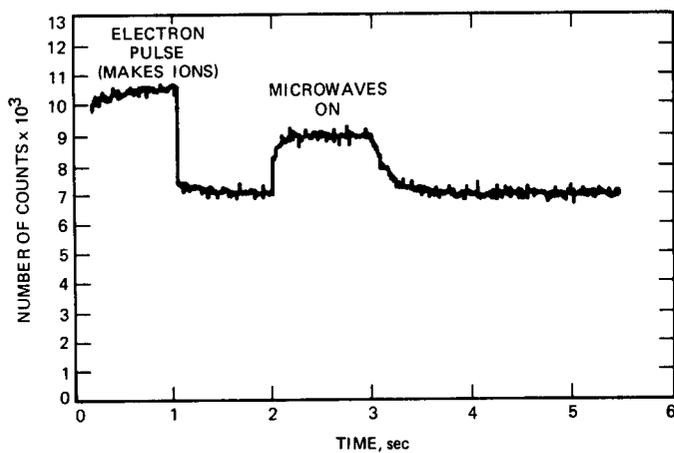


Fig. 6. Light scattered from the ions upon application of microwave radiation (40.5 GHz) is detected together with stray light from electrodes and windows

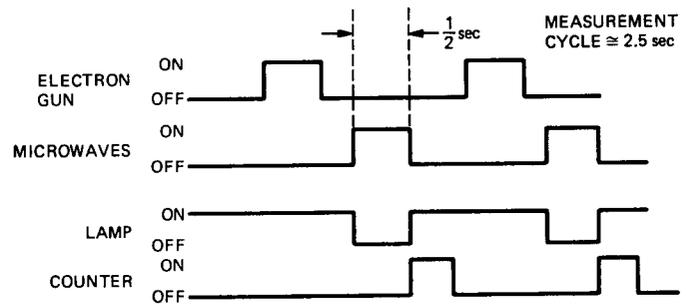


Fig. 7. Sequence of operations used to determine whether the 40.5 GHz is correctly tuned to the atomic resonance (two cycles are shown)

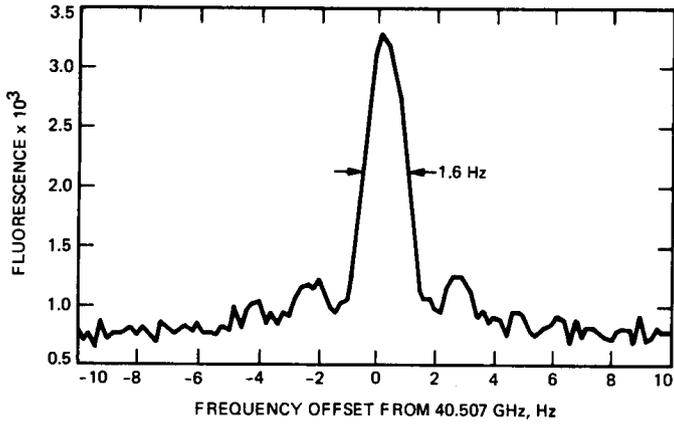


Fig. 8. Atomic fluorescence as the frequency of the 40.5-GHz oscillator is swept through the atomic resonance

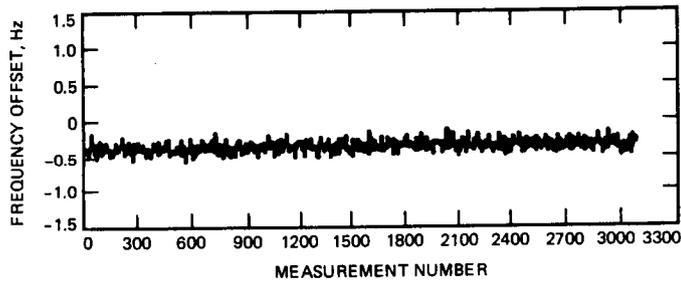


Fig. 9. Frequency deviation of the 40.5-GHz oscillator while servoed to the 1.6-Hz-wide atomic resonance (each measurement lasts 2.5 seconds)

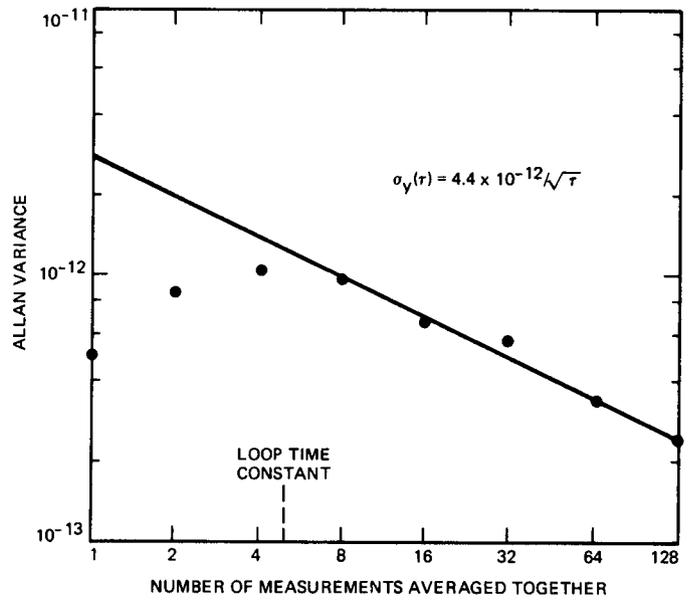


Fig. 10. Allan variance of the sequence of center frequencies shown in Fig. 9; multiply by 2.5 seconds to convert the horizontal axis to time

S3 - 32

128219

TDA Progress Report 42-92

18.

N88-18777

October-December 1987

Using Ridge Regression in Systematic Pointing Error Corrections

C. N. Gujar

Ground Antenna and Facilities Engineering Section

A pointing error model is used in the antenna calibration process. Data from spacecraft or radio star observations are used to determine the parameters in the model. However, the regression variables are not truly independent, displaying a condition known as multicollinearity. Ridge regression, a biased estimation technique, is used to combat the multicollinearity problem. Two data sets pertaining to Voyager 1 spacecraft tracking (days 105 and 106 of 1987) were analyzed using both linear least squares and ridge regression methods. The advantages and limitations of employing the technique are presented. The problem is not yet fully resolved.

I. Introduction

A pointing error model is used in the antenna calibration process to compensate for systematic error sources. Data from spacecraft (s/c) or radio star observations are used to determine the parameters in the model. The model parameters are then used to generate a systematic error correction table for accurately pointing the antenna. The pointing error modeling approach used was originally devised by optical astronomers and subsequently adapted by radio astronomers for RF antennas. The model is based on logical, expected physical behavior of the antenna and has been successfully applied to many radio astronomy facilities: the Bonn 100-m Az-El antenna [1] and the Haystack 37-m Az-El antenna [2]. The complete pointing error model for an antenna is a sum of individual error functions. Table 1 shows the individual error sources and the elevation and cross-elevation (or, depending on the antenna mount, declination and cross-declination) error functions used to develop a systematic error correction table ([1], [2] and [3] give a more in-depth description of the parameters).

When modeling a system, one may select the model purpose to fall into one of three main categories: explanation, variable selection, or prediction. If the model is explanatory, then it represents the y in terms of the x 's and explains how the x 's affect the y . Variable selection techniques should be used when the goal is to determine which variables from a group of variables are important in determining the optimal model for y . This selection of variables could provide the best fit, the simplest form of the model, or both. Prediction, or forecasting, techniques estimate the output, y , at previously unobserved values of inputs, x .

The current pointing error model used in the DSN is of the explanatory type, and the parameters P are determined by performing a linear least squares fit on offset data collected from s/c or radio star observations. Currently, the regressor variables are not truly independent and, rather, display redundant information—a condition known as multicollinearity [4]. Multicollinearity results in limitations on the ability of an ordinary linear least squares fit to provide stable and accurate variables.

It is therefore desirable to study alternate techniques for parameter estimation. Ridge regression is a biased estimation technique for combating the multicollinearity problem. This article reviews the use of the ridge regression technique and demonstrates the advantages and limitations of its uses for systematic error correction development.

II. Review of Regression Analysis

Suppose that, in an experiment, values of the dependent variable y are observed, each corresponding to a particular value of an independent variable x . A straight line representation of the $y = y(x)$ data would have the form

$$y = \beta_0 + \beta_1 x + \epsilon \quad (1)$$

where ϵ is the model error. Equation (1) is a simple linear regression model since it contains a single regressor variable, x , and is linear in x .

The above linear regression of y upon a single variable x can be extended to the multiple linear regression model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \epsilon_i \quad (2)$$

where $i = 1, 2, \dots, n$ ($n \geq k + 1$), ϵ_i is a conceptual random model error assumed to be uncorrelated for each observation (having a zero mean and a constant variance σ^2), x_{ki} are the independent variables (or regressors), y_i are the dependent variables (or response variables) and are the true responses, and β_k are the unknown regression parameters. One equation can be written for each observation, and the error term ϵ allows the model to be an equality. In matrix terms, Eq. (2) becomes

$$\underline{y} = \underline{\beta X} + \underline{\epsilon} \quad (3)$$

Since the regression terms β_k are unknown, let the least squares estimator for these coefficients be b_k . These estimators should satisfy the following equation:

$$\hat{y}_i = b_0 + b_1 x_{i1} + \dots + b_k x_{ik} \quad (4)$$

where \hat{y}_i are the model's estimated (or fitted) value to y_i of Eq. (2). Since Eq. (4) contains only known terms, it does not contain the conceptual terms ϵ_i .

If the initial model was accurate, then the difference between y_i and \hat{y}_i should be small. The difference or residual, r_i , between the actual values and the fitted values is

$$r_i = y_i - \hat{y}_i \quad (5)$$

The method of least squares chooses b_{ik} values so that

$$\sum_{i=1}^n r_i^2$$

is minimized. The estimates satisfy the following matrix equation [4], [5]:

$$\underline{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\underline{y} \quad (6)$$

where \mathbf{X}' is the transpose of \mathbf{X} . When the regressor variables are centered (made dimensionless relative to a mean value), $\mathbf{X}'\mathbf{X}$ is then in correlation form and will be written as $\mathbf{X}^*\mathbf{X}^*$.

III. Multicollinearity

Multicollinearity exists when the regressor variables are empirically correlated, affecting the computation of \underline{b} , which involves the $\mathbf{X}'\mathbf{X}$ matrix. When this situation exists, no conclusions can be drawn as to the individual roles of the variables. If multicollinearity is "severe," then the coefficients may (1) be the wrong size (too large in magnitude); (2) have the wrong sign; or (3) be unstable due to ill-conditioned matrix computations (i.e., small changes in the y 's or x 's lead to large changes in the coefficients). Multicollinearity will also inhibit the ability to predict.

Diagnostics can be performed to evaluate the extent of the multicollinearity problem. Large values in the correlation matrix are one indication of multicollinearity, but this observation only shows pairwise correlations, not correlations that exist between more than two variables. Variance Inflation Factors (VIFs) are another means of identifying multicollinearity. VIFs are the diagonal elements of the inverse of the correlation matrix and represent the inflation that each regression coefficient experiences above the ideal (identity matrix). VIFs are considerably more useful for multicollinearity detection than simple correlation values because they give a direct measure of multicollinearity and tell the user which coefficients are adversely affected and to what extent. As a rule of thumb, VIFs greater than 10 indicate that a severe multicollinearity problem exists. Table 2 gives a sample analysis of a set of conical scanning (conscan) offset data (collected during a Voyager 1 track on the 105th day of 1987) that exhibits a multicollinearity problem. Correlation values of zero mean no correlation and ± 1.0 means full correlation. The VIF data from Table 2 indicates a severe multicollinearity problem.

IV. Ridge Regression

Ordinary least squares methods give unbiased estimates and have the minimum variance of all linear unbiased estimators. However, there is no upper bound on what the variance could be, and the presence of multicollinearity could produce large variances. Ridge regression is a biased estimation technique used to attain a substantial reduction in variance with an increase in the stability of the coefficients. If the correlation matrix is reduced, then the variance

$$\text{var}(\mathbf{b}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \quad (7)$$

is improved and the stability of the coefficients is increased. Ridge regression uses this idea.

Variables x and y in Eq. (2) must first be standardized (centered), making them dimensionless relative to an average value

$$x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{\sqrt{\sum_{i=1}^n (x_{ij} - \bar{x})^2}} \quad (8)$$

$$y_{ij}^* = \frac{y_i - \bar{y}}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (9)$$

where i is the number of points ($i = 1, 2, \dots, n$) and j is the number of parameters ($j = 1, 2, \dots, k$). The new standardized model becomes

$$\underline{y}^* = \mathbf{X}^* \underline{\beta}^* + \underline{\epsilon} \quad (10)$$

and the solution for the least squares estimate \underline{b}_{LS}^* is

$$\underline{b}_{LS}^* = (\mathbf{X}^{*\prime} \mathbf{X}^*)^{-1} \mathbf{X}^* \underline{y}^* \quad (11)$$

where $\mathbf{X}^{*\prime} \mathbf{X}^*$ is the correlation matrix, as stated previously.

The ideal correlation matrix is the identity matrix, \mathbf{I} . If multicollinearity exists, high correlation values exist so the diagonal elements do not dominate and there are large off-diagonal values. To make the correlation matrix values approach the identity matrix, the ridge estimator is introduced:

$$\underline{b}_R^* = (\mathbf{X}^{*\prime} \mathbf{X}^* + k\mathbf{I})^{-1} \mathbf{X}^* \underline{y}^* \quad (12)$$

where \mathbf{I} is the identity matrix and k is a value greater than or equal to zero and is chosen by the user. The term $k\mathbf{I}$ adds a

positive constant to the diagonal elements of the correlation matrix in order to make the diagonal elements dominate. Accordingly, the inverse $(\mathbf{X}^{*\prime} \mathbf{X}^* + k\mathbf{I})^{-1}$ will have smaller elements, alleviating past difficulties created by having large elements on the diagonals of the inverse, like large variances. The term k is often referred to as a "shrinkage parameter" since it "shrinks" the effects of the off-diagonal elements. The ridge estimator \underline{b}_R^* equals the least squares estimator \underline{b}_{LS}^* when $k = 0$. It can also be easily converted back to \underline{b}_R (dimensioned) by a simple transformation.

Ridge regression is called a biased estimation technique since the ridge estimators \underline{b}_R^* are biased. Proper selection of the shrinkage parameter minimizes the negative effect of large bias while maintaining a ridge estimator variance that is significantly less than the least squares estimator. As the shrinkage parameter increases, the bias of the ridge estimator increases and its variance decreases.

A subjective method exists for choosing the shrinkage parameter: the ridge trace. Many different values of k are used to compute $\underline{b}_R^*(k)$, and then each $\underline{b}_R^*(k)$ is plotted versus k . The more unstable the variable is, the faster it drops off and stabilizes. Gradual changes of the variables over k denote stability. The shrinkage parameter k is chosen so that the estimates are stable. As a rule, the smallest value of k where stability of the coefficients first appears is selected [4], [6].

V. Two Case Studies

Two applications of the ridge regression technique on the systematic error correction model were done using Voyager 1 conical scanning (conscan) offset data. The results were compared to fits obtained using an ordinary linear least squares method. The selected parameters for the linear least squares fit were (refer to Table 1) $P_1, P_7, P_8, P_{12}, P_{13}, P_{14}$, and P_{16} . The parameters selected for the ridge regression cases were $P_8, P_{12}, P_{13}, P_{14}$, and P_{16} . Parameters P_1 and P_7 represent constant cross-elevation and elevation offsets, respectively. In the ridge regression process, these two terms were created by determining the cross-elevation and elevation offset biases.

The first data set uses conscan offset data collected on the 105th day of 1987. As demonstrated in Table 2, this data exhibits a high degree of multicollinearity and would probably benefit from the use of ridge regression. Parameters determined using the linear least squares method are listed in column 1 of Table 3. These parameters exhibit the characteristics associated with multicollinearity, one of them consisting of coefficients that are too large in magnitude (they are too large to be realistic or practical). Shrinkage parameters were selected in 0.005 increments and ranged from 0 to 0.10. Fig-

Figure 1 shows the use of the ridge trace for the "best" subjective selection of ridge estimators. Stability seems to be reached at approximately $k = 0.02$. The parameters for this shrinkage parameter are listed in column 2 of Table 3. The coefficients have diminished in value, approaching a more realistic representation. Figure 2 compares the residual fit errors obtained in both the linear least squares method and ridge regression. The residual errors are defined in Eq. (5) as the difference between the actual and the fitted pointing offsets. The signatures for both sets of residual errors are similar, indicating incompleteness in the model itself, but the average residual offset for the ridge regression case is nearly zero, and the standard deviations are similar (approximately 0.9 mdeg).

The above example demonstrated how ridge regression can be used to obtain more realistic parameters and fewer overall fitting errors (average error approaching zero). Multicollinearity also causes the parameters to be unstable. Conscan offset data collected from Voyager 1 tracks on the 105th and 106th days should yield similar results. No changes were made to any part of the antenna mechanical subsystem between these two consecutive tracking sessions (for example, the same *a priori* systematic error correction table and autocollimators were employed in both cases), yet the parameters determined using

the linear least squares fitting method (listed in columns 1 and 3 of Table 3) seem to indicate otherwise. The parameters not only differ in sign, but also differ radically in magnitude. Parameters determined using ridge regression (columns 2 and 4 of Table 3) are in closer agreement in both magnitude (off by a small factor—3 or 4—rather than 10 or 20) and sign, and also yield similar overall fits (same average and standard deviation).

VI. Conclusion

The ridge regression technique was shown to be useful in minimizing the effects of multicollinearity. For the two examples given, it generated stable coefficients for similar sets of data, provided coefficients that were more realistic in magnitude, and gave an overall fit with average residual errors near zero. Although these are good results in terms of coefficient characteristics, the overall fitting results using ridge regression were no better than the linear least squares results since the signatures resulting from the two methods exhibited analogous trends. A technique such as variable selection or prediction may be needed in order to get a more optimal model and a better parameter selection procedure. In any case, the problem of multicollinearity must still be addressed and resolved.

References

- [1] P. Stumpff, "Astronomical Pointing Theory for Radio Telescopes," *Klein-Heibacher Berichte*, vol. 15, Formoldeteshnischon Zentralamt, Darmstadt, West Germany, pp. 431-437, 1972.
- [2] M. L. Meeks, H. A. Ball, and A. B. Hull, "The Pointing Calibration of the Haystack Antenna," *IEEE Transactions on Antennas and Propagation*, vol. AP-16, no. 6, pp. 746-751, November 1968.
- [3] C. N. Guiar, F. L. Lansing, and R. Riggs, "Antenna Pointing Systematic Error Model Derivations," *TDA Progress Report 42-88*, vol. October-December 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 36-46, February 15, 1987.
- [4] R. B. Myers, *Classical and Modern Regression with Applications*, Boston: Duxbury, 1986.
- [5] E. Kreysig, *Advanced Engineering Mathematics*, New York: John Wiley and Sons, 1983.
- [6] A. J. Bush, "Ridge: A Program to Perform Ridge Regression Analysis," *Behavior Research Methods and Instrumentation*, vol. 12, no. 1, pp. 73-74, 1980.

Table 1. Systematic pointing error sources and model terms

Error source	Model function	
	Cross-elevation error	Elevation error
Az collimation	P_1	-
Az encoder fixed offset	$P_2 \cos(\text{el})$	-
Az/el skew	$P_3 \sin(\text{el})$	-
Az axis tilt	$P_4 \sin(\text{el}) \cos(\text{az})$	$-P_4 \sin(\text{az})$
Az axis tilt	$P_5 \sin(\text{el}) \sin(\text{az})$	$P_5 \cos(\text{az})$
El encoder fixed offset	-	P_7
Gravitational flexure	-	$P_8 \cos(\text{el})$
Residual refraction	-	$P_9 \cot(\text{el})$
Az encoder scale error	$P_{10} (\text{az}/360) \cos(\text{el})$	-
	Cross-declination error	Declination error
HA/dec axis skew	$-P_{11} \sin(\text{dec})$	-
HA axis tilt	$P_{12} \sin(\text{HA}) \sin(\text{dec})$	$P_{12} \cos(\text{HA})$
HA axis tilt	$-P_{13} \cos(\text{HA}) \sin(\text{dec})$	$P_{13} \sin(\text{HA})$
HA feed offset	$-P_{14}$	-
Gravitational flexure	$P_{15} \cos(p) \cos(\text{el})$	$-P_{15} \sin(p) \cos(\text{el})$
Declination feed offset	-	P_{16}
Gravitational flexure	$P_{17} \sin(p) \cos(\text{el})$	-
Gravitational flexure	-	$-P_{18} \cos(p) \cos(\text{el})$
Gravitational flexure	$-P_{19} \sin(\text{el})$	-
Gravitational flexure	-	$P_{20} \sin(\text{el})$
HA encoder bias	$P_{21} \cos(\text{dec})$	-

Note: (1) Uppercase P refers to parameter value; lowercase p refers to paralectic angle.
(2) Az = azimuth angle; el = elevation angle; dec = declination angle; HA = hour angle.

Table 2. Sample correlation matrix and variance inflation factors (for Voyager 1 conscan offset data from 105th day of 1987)

Variable	Correlation matrix					VIF
	8	12	13	14	16	
8	1.0000	0.8653	-0.9911	-0.9954	0.9937	747.6
12	0.8653	1.0000	-0.8624	-0.8690	0.8994	103.9
13	-0.9911	-0.8624	1.0000	0.9981	-0.9959	2783.1
14	-0.9954	-0.8690	0.9981	1.0000	-0.9966	929.0
16	0.9937	0.8994	-0.9959	-0.9966	1.0000	4377.6

Table 3. Model parameters for two Voyager 1 conscan offset data sets (105th and 106th days of 1987) using linear least squares and ridge regression (units are in millidegrees)

Parameter (P)	Day 105		Day 106	
	Linear least squares (1)	Ridge regression (2)	Linear least squares (3)	Ridge regression (4)
1	-451.93	24.08*	-0.55	18.68*
7	-141.18	29.58*	17.86	17.13*
8	-271.78	-13.45	-27.22	-29.11
12	-114.05	4.28	1.40	1.34
13	240.03	-4.10	-0.65	-17.24
14	-557.56	4.85	6.19	15.26
16	103.29	0.25	0.49	0.89

*Ridge regression parameters P_1 and P_7 are created by determining the cross-elevation and elevation biases.

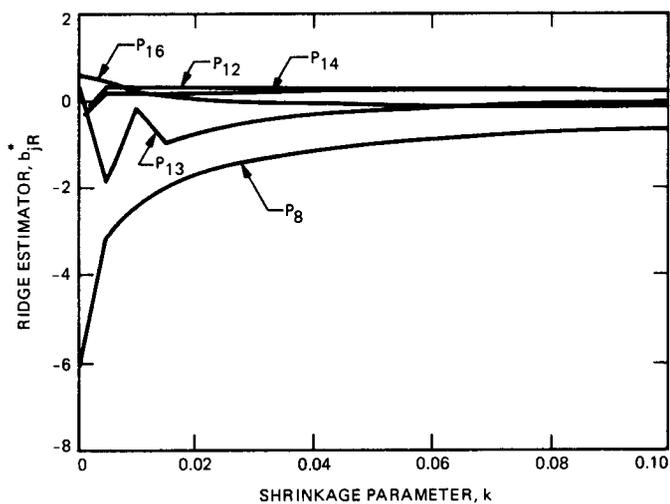


Fig. 1. Ridge trace for conscan offset data from a Voyager 1 track collected on 105th day of 1987 shows parameters reaching stability at approximately $k = 0.02$

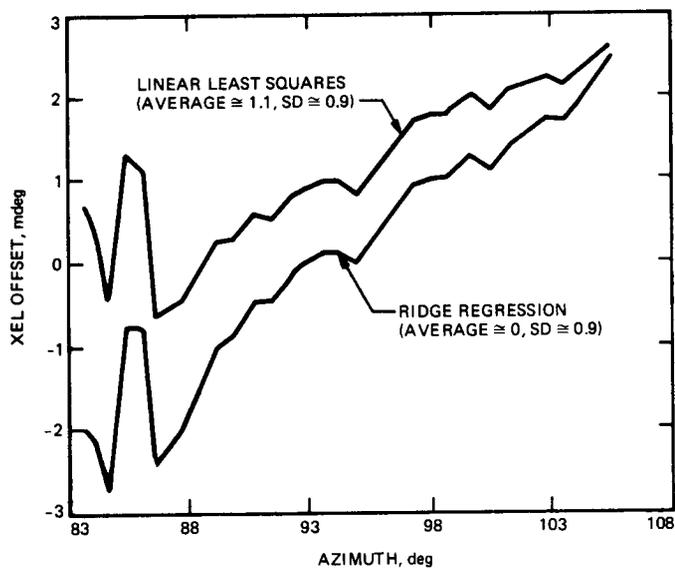


Fig. 2. Comparison of residual pointing errors (difference between actual and fitted offsets) using the linear least squares and ridge regression for data collected on 105th day of 1987

An Optimal Structure for a 34-Meter Millimeter-Wave Center-Fed BWG Antenna: The "Cross-Box" Concept

K. L. Chuang

Ground Antenna and Facilities Engineering Section

This article presents a new approach to the design of the planned NASA/JPL 34-m elevation-over-azimuth (Az-El) antenna structure at the Venus site (DSS-13). The new antenna structural configuration accommodates a large (2.44-m) beam waveguide (BWG) tube centrally routed through the reflector-alidade structure, a unique elevation wheel design, and an optimal structural geometry. The new design encompasses a "Cross-Box" elevation wheel-reflector base substructure that preserves homology while satisfying many constraints, such as structure weight, surface tolerance, stresses, natural frequency, and various functional constraints. The functional requirements are set to ensure that microwave performance at millimeter wavelengths is adequate.

The new Cross-Box configuration was modeled, optimized, and found to satisfy all DSN HEF baseline antenna specifications. In addition, the new structure design was conceptualized and analyzed with an emphasis on preserving the structure envelope and keeping modifications relative to the HEF antennas to a minimum, thus enabling the transferability of the BWG technology for future retrofitting. Good performance results were obtained.

I. Introduction

The DSN is planning to build a new R&D antenna at the Venus Deep Space Tracking Station (DSS-13) at Goldstone, California [1]. The proposed R&D antenna is intended to perform as a test bed for the development of advanced telecommunication technologies, among which are the incorporation of (1) beam waveguide optics, (2) Ka-band (32 GHz) components, and (3) high gain/noise temperature capability in the millimeter wavelength range.

The inclusion of a BWG system is viewed as an item of high research priority [2] for the existing baseline design of the 34-m HEF network (Figs. 1 and 2). It is also proposed that any new designs be capable of retrofit to the existing network of antennas at low cost and that all future DSN antennas make use of the BWG optics.

The modeled Cross-Box configuration presented in this article is one of several options that satisfy the above conditions while introducing minimal changes in the antenna geometry

and tipping structure weight relative to present 34-m DSN HEF antennas. The new design also maintains the major features of existing drive systems for the Az-El mount, the alidade structure, and the azimuth wheel and track. Proper design of the elevation wheel and support structure (ELWH) obviates the need to redesign the main reflector and its backup truss (REFL). Thus, optimal configuration of the antenna structure primarily entails the design of the substructure that transfers loads from the main reflector to the elevation bearings only (Fig. 2)—herein called substructure ELWH. This article describes the design approach and finite element model formation.

II. Design Statement

The "Cross-Box" design was modeled as a truss-type structure with all joints (nodes) modeled as pinned joints. Connectivity between nodes can be achieved with bars/rods (one-dimensional elements) or with plates (triangular or quadrilateral). All elements are assumed to be non-bending members. Altogether, there are approximately 3900 members and 1200 nodes in the whole antenna structure above the azimuth track. For optimization purposes, the members are judiciously grouped to maintain antenna symmetries, resulting in 207 groups of design variables for the problem. The design variables comprise bar areas and plate thicknesses.

Determination of the sizes of the members is the eventual problem that must be solved. Optimization schemes to obtain explicit solutions are available in the literature for structures in general [3]–[10] and for antennas in particular [11]. Not as easily resolved are problems of establishing optimal configurations, especially those in which path obstructions are encountered in the structural geometry. The major difficulty in designing for the present antenna is of this nature. Discussion on the course taken to overcome it will be emphasized. The complexities of the real antenna structure geometry and component fabrication do not readily lend themselves to simple solutions. Shape optimization schemes may help but they are mostly of the perturbation type [12]–[18]. Effectively, this amounts to perturbation of nodes in a finite element model. However, attempting to weave a tube of eight-foot diameter within the antenna structure demands more than minor adjustments in geometry. By and large, when solution strategies exist, they are usually problem-specific [12], [19]–[23]. Variational methods [23] are not applicable to structures of highly noncontiguous domain and therefore are of no value in the present problem. Although the layout theory propounded in [19]–[22] is applicable to 1- and 2-dimensional problems, it is extremely difficult to apply to 3-dimensional problems. Nevertheless, it is unlikely that any configuration under optimization will evolve into a dramatically different geometry

without creating other problems, such as overlapping/intersecting members.

When a viable configuration is found, the problem is "reduced" to one that is amenable to classical treatment of feasibility, Lagrange multipliers, Kuhn–Tucker conditions, etc. For larger problems, indirect methods are used whereby the notion of active and passive constraints is introduced. Optimality criteria methods [4]–[11] are based on such principles. Optimal member sizes for the Cross-Box antenna were obtained through the use of NASA/JPL–IDEAS programs [11], which employ a version of the optimality criteria methods [9]. The Cross-Box design antenna has more than 11,700 degrees of freedom, which is common for antennas of this size. Stiffness matrix decomposition time is significant and is compounded when designing the parabolic surface of the antenna. Attempts have to be made to best-fit thousands of points of the reflector surface to a paraboloid at different antenna elevations [26].

Another problem arises from the degree of static indeterminacy of the structure. The more statically indeterminate the structure is, the harder it is to predict the inputs for the next iteration [9]. The methodology becomes sensitive to step-size/move limits chosen in the iterations. Not only does it take longer to converge on a solution but it is also possible for the algorithm to fail to yield a reasonable solution. In addition, other solution-strategy-dependent features may affect computational efficiency. Perhaps the most direct impositions come from side constraints. They increase computational time, if nothing else. At worst, these constraints might be such that the problem will have no feasible solution. The following are the dominant constraints that affect the design of the Cross-Box antenna:

- (1) Maximum allowable member stresses (yield and buckling).
- (2) Upper and lower bounds of member sizes.
- (3) Maximum allowable displacements at worst elevation angle:
 - (a) Surface distortion (rms) from a best-fit paraboloid due to gravity loading.
 - (b) Surface distortion (rms) from a best-fit paraboloid due to worst case wind (120 degrees elevation and 0 degrees azimuth) loading.
- (4) Boresight error due to worst case wind (0 degrees elevation and 120 degrees azimuth) loading.
- (5) Weight of tipping structure limited to 220 ± 10 percent kips.
- (6) Lowest natural frequency.

- (7) Survivability of antenna in stow position (at zenith) for a 100 mph wind.

Other less quantifiable constraints on conceptualizing the configuration are as follows:

- (1) The structure shall be such that an elevation range of 6 to 90 degrees is possible.
- (2) The structure shall be designed around the space required by the BWG path (as dictated by BWG optics) within the antenna structural domain.

The structure should also have the following attributes:

- (1) A simple configuration geometry.
- (2) Minimal member connections/joints.
- (3) Low-cost fabrication techniques.
- (4) Duplicability of substructures and symmetries.
- (5) Duplicability and retrofittability of the configuration to the existing DSN HEF antenna (this also implies that the design can adopt components from existing antennas).

These attributes will directly translate into reduced cost impact on the design.

III. Model Formation

To reach an optimum configuration for the antenna structure, steps were taken following the schematic of Fig. 3. The blocks indicate the end products while the arrow paths denote actions. In general, the design optimization strategy involves the following procedures:

- (1) Conceive modularization/substructuring.
- (2) Design and optimize each substructure.
- (3) Synthesize the antenna by assembling individually optimized substructures, followed by optimization of the antenna as a whole using results obtained in (2) as the best input prediction.

The above idealization assumes that the synthesis of optimum substructures will yield an optimum global structure through proper choice of modularization. Each procedure will now be described.

A. Modularization

In designing the antenna structure, some modularization is advantageous because it

- (1) Allows the problem to be more easily managed.
- (2) Allows the problem modules to be tackled separately and simultaneously—and thus more efficiently.
- (3) Allows individual modules to be solved on microcomputers, thus requiring less time on a “large” program run on a “large” computer.
- (4) Allows the convenience of making changes to, and testing of, individual modules without having to run the whole program/antenna structure.
- (5) Allows different degrees of difficulty encountered in different substructures to be treated by specialized programs. (For example, the NASTRAN program analyzes structures with bending members but NASA/JPL-IDEAS does not. On the other hand, the latter performs structural optimization and also surface best-fitting, which the former does not.)
- (6) Allows reduced mass storage.

The 34-m antenna at hand is divided into three substructures: (1) the main reflector and its backup truss, subreflector, and mount (REFL); (2) the elevation wheel structural assembly (ELWH); and (3) the alidade (ALD). They are distinguished by their distinct structural functions as shown in Fig. 2. The REFL module upholds a parabolic surface, allowing minimal surface distortion from some parabolic (shaped) profile, maintaining homology [24], and promoting symmetrical displacement of surface points. The ELWH module transfers loads (gravity and wind) from the REFL module to the ALD module and simultaneously counterbalances the REFL about the elevation axis. When the antenna is at zenith position, the ELWH module ideally provides a plane surface on which the REFL can be placed. When the antenna is at the horizon position, the ELWH, with the counterweight that it carries, balances the REFL, thus relieving the elevation drives from load bearing. The elevation drive is the only other point besides the two elevation bearings at which there is contact between the tipping structure and the alidade.

The third substructure, ALD, bears all the loads from the tipping structure transmitted through the elevation bearings. It provides the azimuth range for the antenna driven on a track. It also provides a mount for the elevation drive. Proper design of the antenna adheres to the principles described in [25].

B. Substructure Design and Optimization

The REFL module design of the NASA/JPL HEF antennas was kept intact: it consists of a parabolic reflector with radial ribs, hoops, and supporting truss. It is a symmetric structure as shown in Fig. 4. For a Cassegrain-type antenna, the secondary reflector and its quadripod mount are included in the REFL

design. The unique features of the new Cross-Box antenna come from the ELWH module with its homology features [24], [25].

The conceptualization of the ELWH design is based on the following observations and reasoning:

- (1) Eight points (on the circumference of a circle) forming the vertices of an octagon are selected at the base of the REFL substructure (Fig. 5). These are attachment points of the REFL to the ELWH enabling uniform displacement of the REFL under symmetric zenith loading. An octagonal truss system is conceived to be part of the ELWH where these attachments are enabled (Fig. 6).
- (2) Loads on the tipping structure must eventually be borne by the two elevation bearings. This implies that connections must be made from eight points to two points. However, this cannot be done directly if uniform deflection of the REFL is to be maintained under uniform zenith loading. Reactions at the eight points cannot be identical unless both points lie along the zenith axis. This is not allowed, however, because a BWG system is to be centrally routed. The solution to this problem is achieved by transferring the loading first to four points (four corners of a square) and then to two points.
- (3) The elevation wheel (bullgear) lies in the Y-Z plane perpendicular to the elevation axis (a line joining the two elevation bearings). It must be incorporated into the ELWH substructure and must exert even loading on the four points in (2). An elevation "axle" and the elevation wheel suggest that a "+" structural form is needed in the ELWH substructure.

Although an elevation "axle" is needed, BWG optics requirements demand that an eight-foot-diameter (2.44 m) path along the elevation axis be devoid of any structure. This means that a box must be built up to assume an elevation axle. Furthermore, to satisfy the criteria in (3), a cross-box must be constructed. This gives rise to the "Cross-Box" design. Figure 7 shows the position of the Cross-Box in relation to the octagonal truss. The "square" mentioned in (2) is dictated by nodes 9, 10, 11, and 12. Bracing members are not shown in Fig. 7. Elevation bearings are a radius distance from the center of the octagon and two feet below the lower edge of the truss. The box is tapered, with the thickest possible section at the central portion. This design feature ensures that flexural deflection of the "axle" is minimized. Figure 8 shows the complete envelope of the cross-box and the BWG path. Bracings for all faces (except 9-10-11-12-9, 25-26-27-28-31-32-21-22-25, and the four faces typified by 31-30-46-31) are not shown. Bar elements are grouped according to their symmetry about the X-Z

or Y-Z plane in conjunction with their positions. The cross-sectional thickness (of the tapered portion) of the x-direction leg of the Cross-Box is limited by the vertical section of the BWG path.¹ This design enables an elevation range of 6 to 90 degrees. Also, the elevation wheel (bullgear) is attached to the 8 points: 41, 23, 24, 42, 43, 29, 30, and 44. It has symmetry about both the X-Z and Y-Z planes and has a design similar to that of the DSN HEF antennas.

The transfer of loads from the REFL to the ELWH is done by connecting bar members from the octagonal truss to the cross-box. Figure 9a shows a plan view of how the connection is done. The "X" bracings shown are conceived to restrain relative rotation about the reflector local Z-axis between the cross-box and the octagonal truss. Note that the four points on the X-Z plane are equally loaded for Z-loading. Members symmetrical about the X-axis and those about the Y-axis are grouped differently to allow homology of structure. This allows for reduced RMS distortion due to Y-direction loading which creates an antisymmetric displacement pattern. Next, rigidity of the box structure is ensured by bracings as shown in Fig. 9b. The "X" bracing indicated by the broken lines in the center indicates that it is at the bottom face; the top face is open for the BWG path. Torsional rigidity of the cross-box about the local Z-axis is established by the bracing shown in the "eye view." The broken circle shows the vertical portion of the BWG tube.¹

The REFL substructure is optimized by restraining the eight attachment points from translation in all directions. The Rigging-Angle Method [26] was used to determine the worst root-mean-square distortion and pointing error (angle) of the best-fit parabolic surface. (The Rigging Angle Method determines the elevation angle for which, if a perfect paraboloid is designed, the worst distortion rms from a paraboloid will occur identically at the zenith and horizon looks of the antenna.) The following results were achieved:

- (1) Surface rms distortion from a best-fit paraboloid due to gravity loading alone is 0.005 inch.
- (2) Surface rms distortion from a best-fit paraboloid due to wind at 30 mph (120 degrees elevation and 0 degrees yaw) is 0.009 inch.
- (3) Boresight error due to a 30 mph wind (0 degrees elevation and 120 degrees yaw) is 0.003 degree.

These results depict the individual effect of the loads applied independently. Optimal sizing of the members has resulted in satisfying specifications as shown in Table 1.

¹Actually, this segment of the BWG tube is not vertical but tilted at an angle of 13.5 degrees to the vertical in the Y-Z plane.

An ALD substructure design can be constructed with little modification from the baseline antennas. Construction of ALD poses no difficulties since the only loadings are applied forces and moments at the elevation bearings. Performance characteristics similar to those of the DSN antennas were obtained with little modification in the design.

C. Substructure Synthesis and Optimization

Figure 10 shows a sketch of an assembled cross-box antenna. Note how the elevation wheel is attached to the components discussed in Figs. 7 and 8. Also seen are the tapered legs of the cross-box and the octagonal truss. The baseline antenna's structural envelope is preserved. Hence, the condition of the fewest changes possible to antenna subsystems is observed. Figures 10 and 11 show essentially the synthesized structure and its accommodation of a center-fed BWG system as dictated by microwave optics.

Optimization was performed on the antenna structural model combining REFL and ELWH. Member sizes in the REFL were not allowed to change (thus allowing direct usage of the DSS 15 design). Results obtained after optimization on member sizes satisfy all specifications. Table 1 lists performance indices achieved. Also cited are the DSN HEF antenna specifications.

IV. Summary

This article reveals the new Cross-Box design concept proposed for the planned 34-m-diameter development antenna at the Venus site. The proposed design has the accommodability of a large beam waveguide (2.4-m) system for Ka-band operability and retrofittability to the 34-m high efficiency antennas. The Cross-Box antenna is optimal in both structural configuration and member size, satisfying many functional constraints. Observance of structural compatibility with the 34-m antennas allows transferability of technologies.

Acknowledgment

The author would like to thank Smoot Katow, John Cucchissi, and Dr. Roy Levy for their helpful suggestions. Dr. Fikry Lansing offered much encouragement and support throughout this work.

References

- [1] J. G. Smith, "Proposed Upgrade of the Deep Space Network Research and Development Station," *TDA Progress Report 42-88*, vol. October-December 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 158-163, February 15, 1987.
- [2] R. C. Clauss and J. G. Smith, "Beam Waveguides in the Deep Space Network," *TDA Progress Report 42-88*, vol. October-December 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 174-182, February 15, 1987.
- [3] L. Schmit and H. Miura, "Approximation Concepts for Efficient Structural Synthesis," NASA CR-2552, March 1976.
- [4] N. Khot, L. Berke, and V. Venkayya, "Comparison of Optimality Criteria Algorithms for Minimum Weight Design of Structures," *AIAA Journal*, vol. 17, pp. 182-190, February 1979.
- [5] S. Segenreich, N. Zouain, and J. Herskovits, "An Optimality Criteria Method Based on Slack Variables Concept for Large Scale Structural Optimization," in *Proceedings of the Symposium on Applications of Computer Methods in Engineering*, School of Engineering, University of Southern California, pp. 563-572, August 1977.
- [6] L. Berke and N. Khot, "Use of Optimality Criteria Methods for Large Scale Systems," in *Proceedings of the AGARD Symposium on Structural Optimization*, AGARD Lecture Series No. 70, pp. 1-29, October 1974.
- [7] C. Fleury, "An Efficient Optimality Criteria Approach to the Minimum Weight Design of Elastic Structures," *Computers and Structures*, vol. 2, pp. 163-173, 1980.
- [8] V. B. Venkayya, N. S. Khot, and L. Berke, "Application of Optimality Criteria Approaches on Automated Design of Large Practical Structures," in *Proceedings of the Second Symposium on Structural Optimization*, Milan, Italy, pp. 3.1-3.19, 1973.
- [9] R. Levy and W. Parzynski, "Optimality Criteria Solution Strategies in Multiple-Constraint Design Optimization," *AIAA Journal*, vol. 20, no. 5, pp. 708-715, April 13, 1981.
- [10] E. J. Haug and J. S. Aurora, *Applied Optimal Design*, New York: John Wiley, 1979.
- [11] R. Levy and R. J. Melosh, "Computer Design of Antenna Reflectors," in *Proceedings of ASCE 99*, pp. 2269-2285, November 1973.
- [12] K. Wasserman, "Three-Dimensional Shape Optimization of Arch Dams with Prescribed Shape Functions," *Journal of Structural Mechanics*, vol. 11, no. 4, pp. 465-489, 1983-84.
- [13] A. Francavilla, C. V. Ramakrishnan, and O. C. Zienkiewicz, "Optimization of Shape to Minimize Stress Concentration," *J. of Strain Analysis*, vol. 10, no. 2, pp. 63-70, 1975.
- [14] O. C. Zienkiewicz and J. S. Campbell, "Shape Optimization and Sequential Linear Programming," in *Optimum Structural Design*, R. H. Gallagher and O. C. Zienkiewicz (eds.), New York: John Wiley, pp. 109-126, 1973.
- [15] M. E. Botkin and J. A. Bennett, "Shape Optimization of Three-Dimensional Folded Plate Structures," presented at the 25th Structures, Structural Dynamics, and Materials (SDM) Conference, Palm Springs, California, 1984.

- [16] S. D. Rajan and A. D. Belegundu, "A Shape Optimization Approach Using Fictitious Loads as Design Variables," presented at the 28th Structures, Structural Dynamics, and Materials (SDM) Conference, Monterey, California, 1987.
- [17] "The Optimum Shape: Automated Structural Design," an international symposium sponsored by General Motors Research Laboratories, Warren, Michigan, October 1985.
- [18] E. J. Haug, K. K. Choi, J. W. Hou, and Y. M. Yoo, "Shape Optimal Design of Elastic Structures," presented at the International Symposium on Optimum Structural Design, University of Arizona, Tucson, Arizona, 1981.
- [19] G. I. N. Rozvany, "A General Theory of Optimal Structural Layouts," presented at the International Symposium on Optimum Structural Design, University of Arizona, Tucson, Arizona, 1981.
- [20] G. I. N. Rozvany and R. D. Hill, "A Computer Algorithm for Deriving Analytically and Plotting Structural Layout," *Comp. and Struct.*, vol. 10, pp. 295-300, 1979.
- [21] G. I. N. Rozvany, C. M. Wang, and M. Dow, "Prager-Structures: Arch-Grids and Cable Networks of Optimal Layout," *Comp. Meth. Appl. Mech. Engrg.*, vol. 31, 1982.
- [22] W. Prager and J. E. Taylor, "Problems of Optimal Structural Design," *Journal of Appl. Mech.*, vol. 35, no. 1, pp. 102-106, 1968.
- [23] J. P. Zolesio, "Domain Variational Formulation for Free Boundary Problems," *Optimization of Distributed Parameter Structures*, E. J. Haug and J. Cea (eds.), Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands, pp. 1165-1208, 1981.
- [24] S. Von Hoerner, "Homologous Deformations of Tilttable Telescopes," in *Proceedings of the ASCE*, pp. 461-485, October 1967.
- [25] S. Von Hoerner, "Design of Large Steerable Antennas," *The Astronomical Journal*, vol. 72, no. 1, February 1967.
- [26] R. Levy, "A Method for Selecting Antenna Rigging Angles to Improve Performance," Deep Space Network Space Programs Summary 37-65, vol. II, Jet Propulsion Laboratory, Pasadena, California, pp. 72-76, July 1970.

Table 1. Achieved performance indices

Antenna type	Weight (kips)	RMS distortion due to gravity alone (inches)	RMS distortion due to 30 mph wind: worst case (inches)	Boresight error due to 30 mph wind: worst case (mdeg)
Cross-box antenna	244	0.009	0.009	10
DSN HEF antenna	217	0.015	0.019	13*

*This figure includes the alidade. A figure for the tipping structure alone was not available. From this estimate, the alidade contribution is approximately 4 millidegrees.

ORIGINAL PAGE IS
OF POOR QUALITY

~~ORIGINAL PAGE IS
OF POOR QUALITY~~

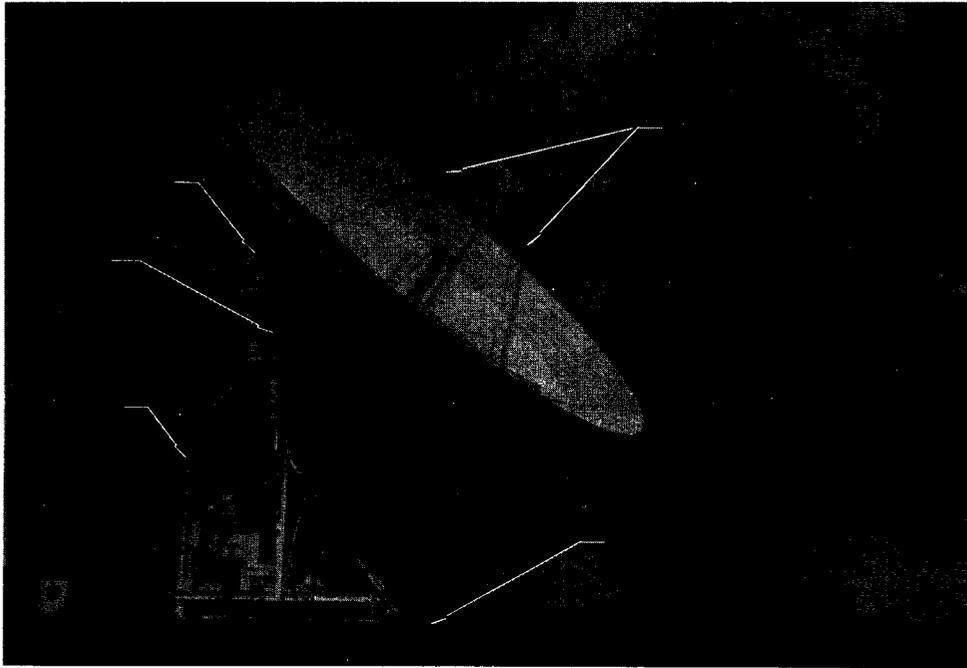


Fig. 1. NASA DSN HEF antenna at Goldstone, California (two identical antennas are located in Spain and Australia)

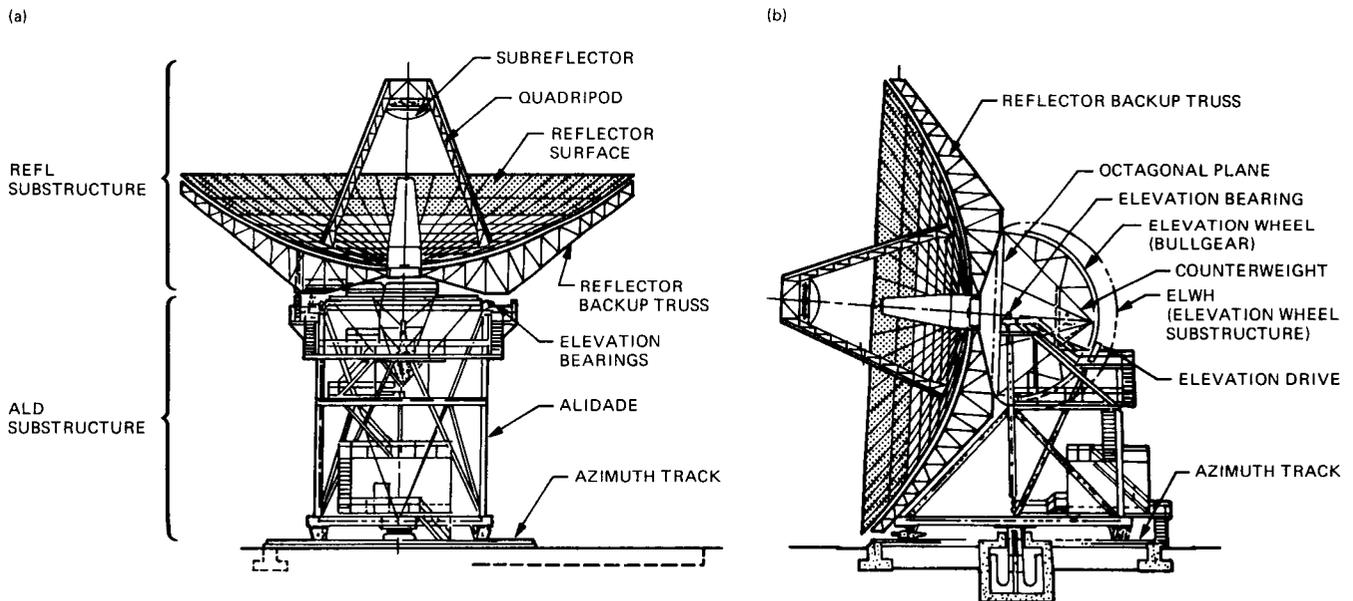


Fig. 2. NASA DSN HEF 34-m antenna (a) at zenith position and (b) at horizontal position

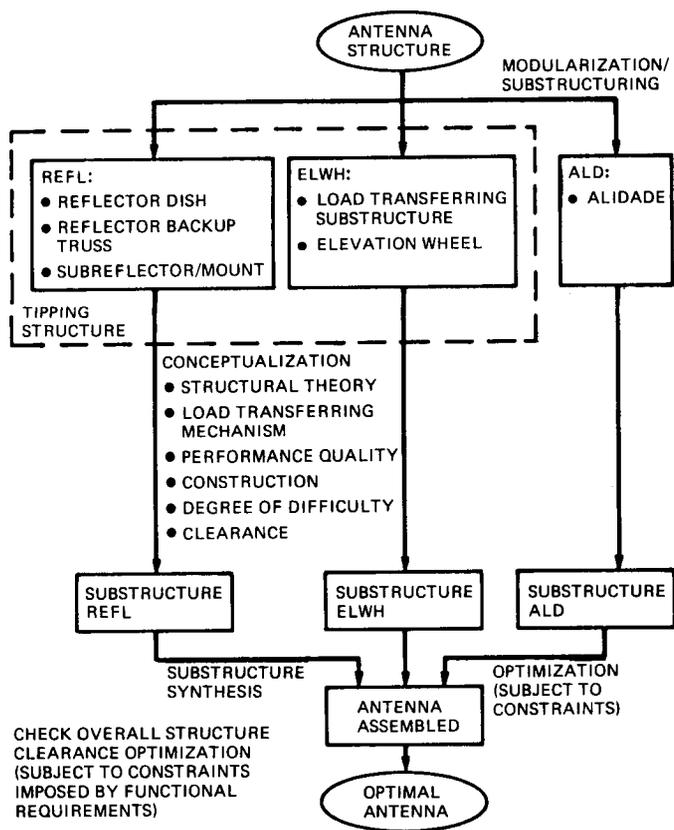


Fig. 3. Schematics showing approach to antenna configuration

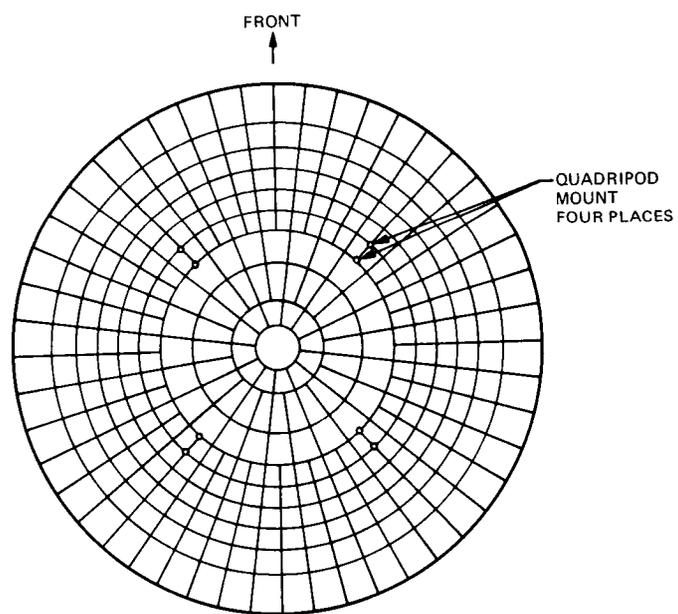


Fig. 4. Front surface of the 34-m HEF antenna main reflector

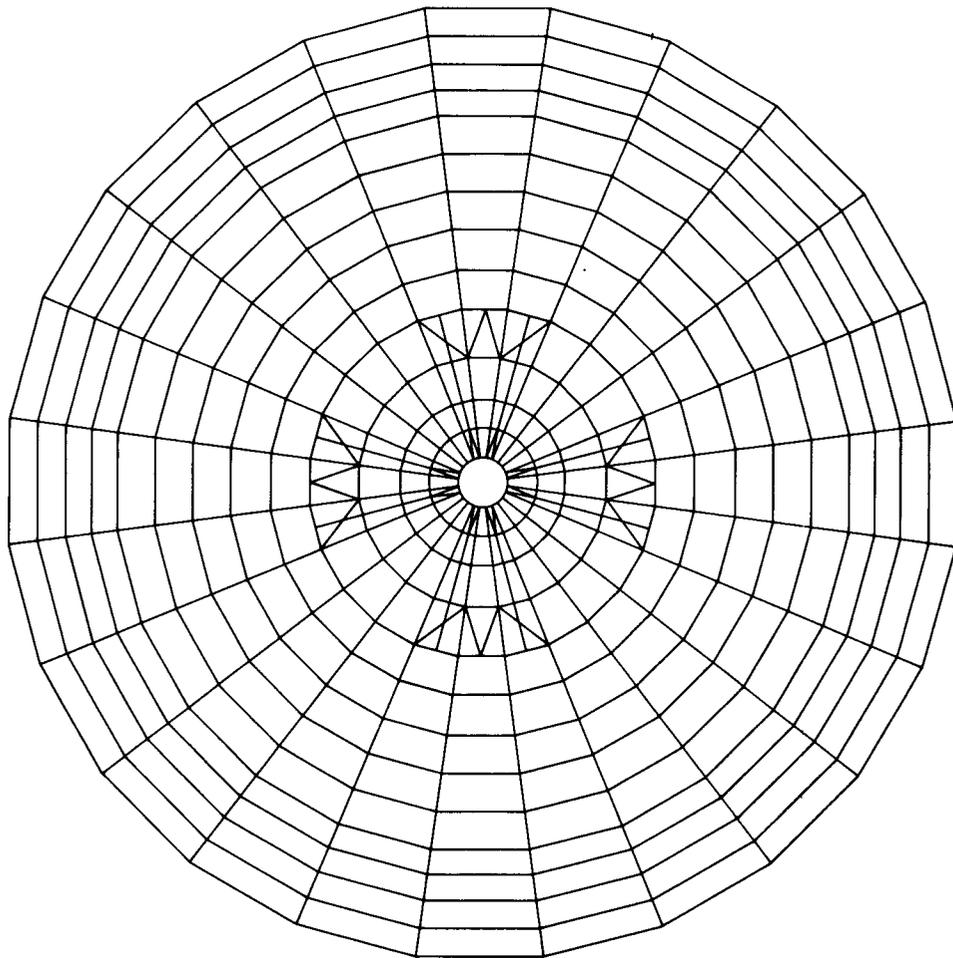


Fig. 5. Finite element model of the reflector (back) showing eight points where attachment of REFL and ELWH will occur

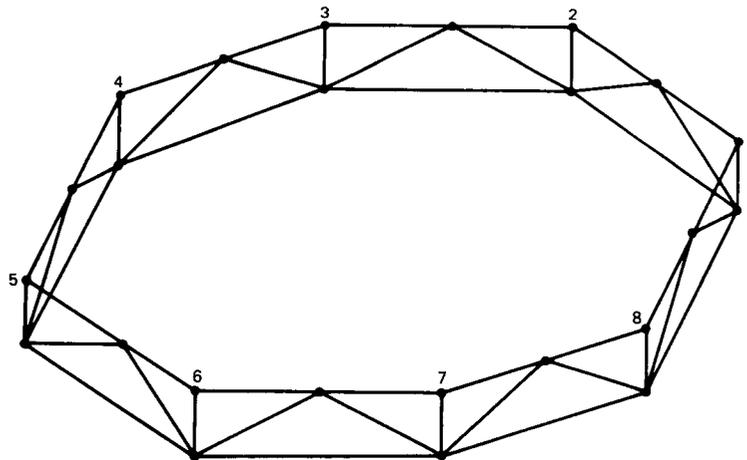


Fig. 6. The octagonal truss

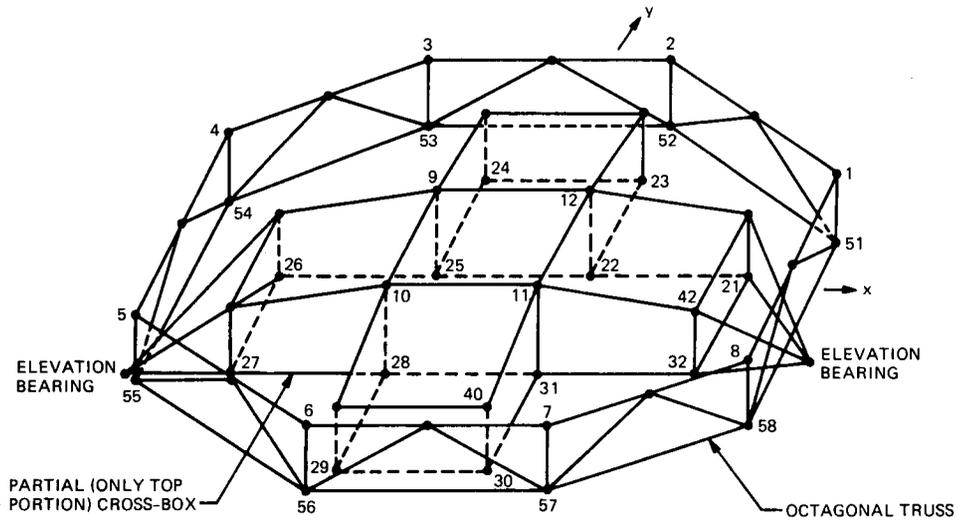


Fig. 7. The cross-box concept showing relative position of cross-box to octagonal truss

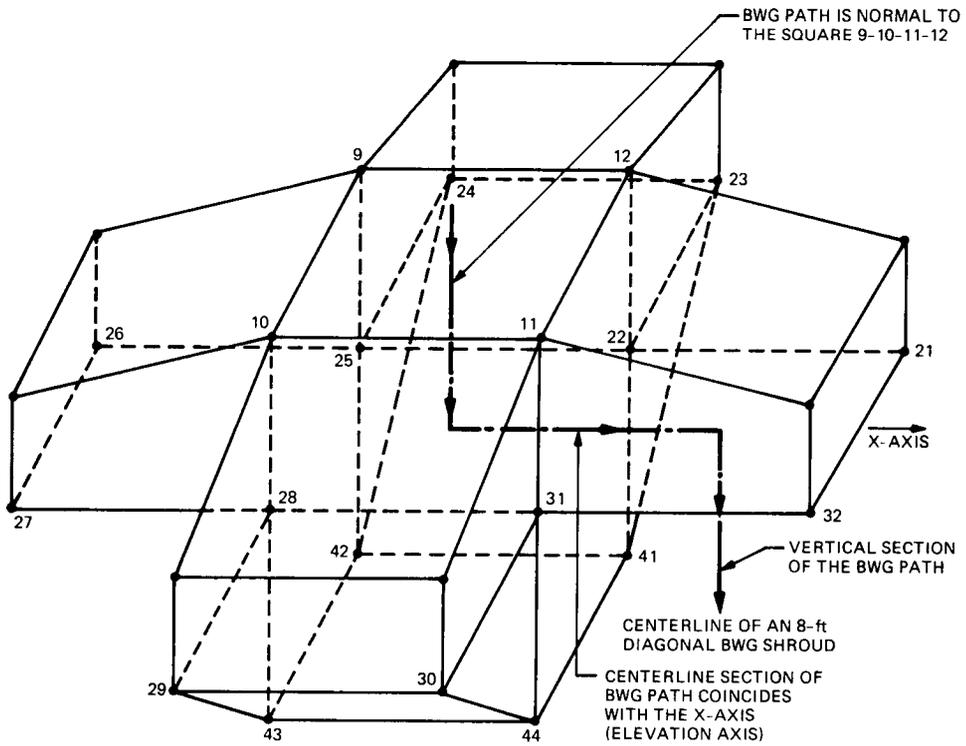


Fig. 8. Center piece (cross-box) showing BWG path

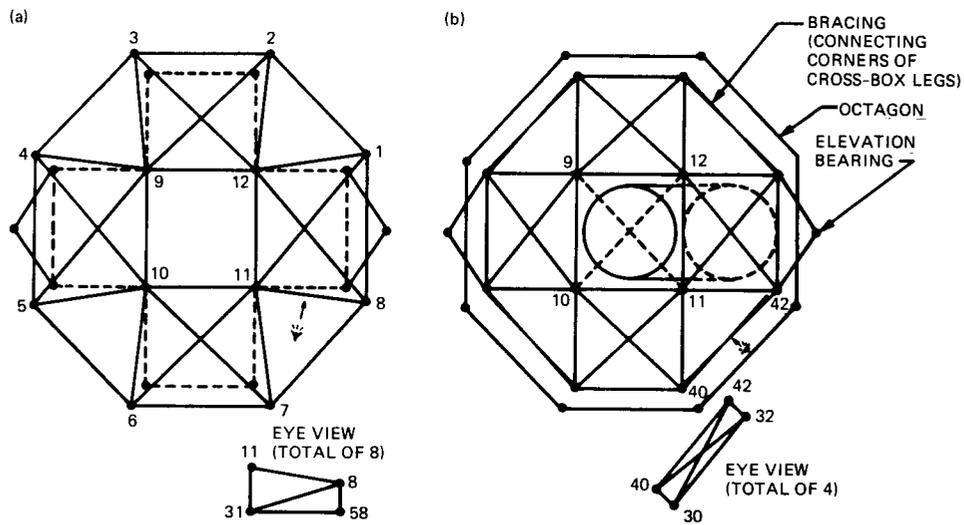


Fig. 9. Connection of bar members from octagonal truss to cross-box: (a) view emphasizing connecting links; (b) view emphasizing cross-box bracing

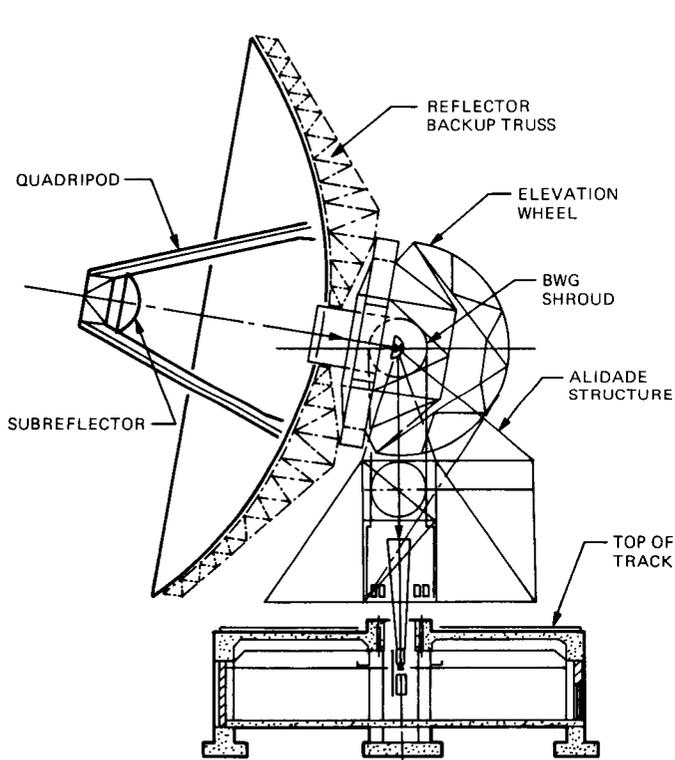


Fig. 10. Approximate relative position and dimension of BWG path to the antenna (Y-Z plane)

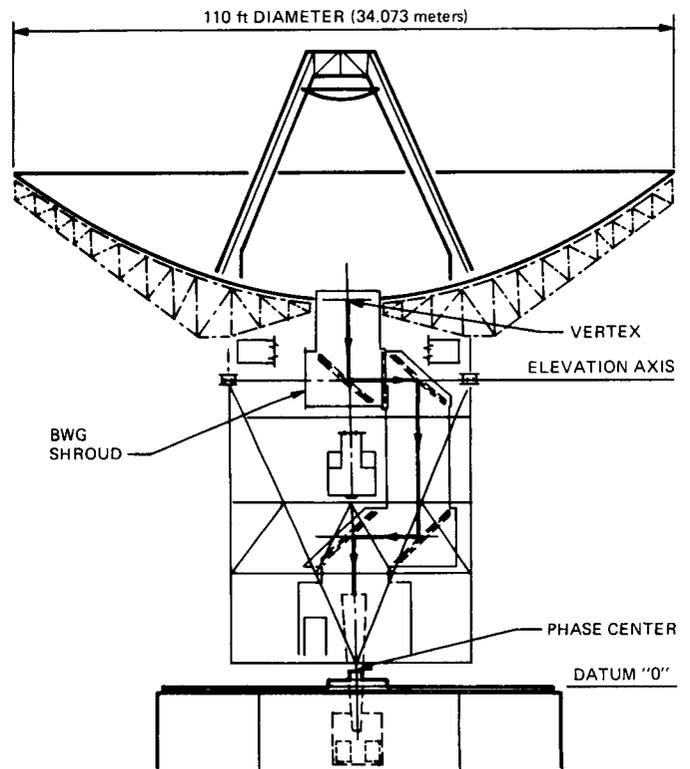


Fig. 11. Approximate relative position and dimension of BWG path to the antenna (X-Z plane)

5536

N88-18779 ;

128221

TDA Progress Report 42-92

October-December 1987

38

A 2.3-GHz Maser at Usuda, Japan, for TDRSS-Orbiting VLBI Experiment

R. B. Quinn

Radio Frequency and Microwave Subsystems Section

A 2.3-GHz traveling-wave maser/closed-cycle refrigerator (TWM/CCR) that is used in the DSN was installed and successfully operated on the 64-m antenna at Usuda, Japan. The TWM/CCR supported the first very long baseline interferometry (VLBI) experiment to use an orbiting spacecraft as one of the receiving antennas. The experiment required a 15-K receiving system over a 2271- to 2285-MHz bandwidth. The maser installation was made during June 1986, and successful VLBI measurements were made during July and August 1986 and again in January 1987.

I. Introduction

A request was made to the Radio Frequency and Microwave Subsystems Section at JPL to install and operate a traveling-wave maser/closed-cycle refrigerator (TWM/CCR) on the 64-meter antenna at the Usuda Deep Space Center (UDSC) near Usuda, Japan. The 15-K receiving system was required for the conceptual demonstration of a very long baseline interferometry (VLBI) experiment that uses an orbiting receiving station at one end of the baseline [1], [2]. The TWM/CCR was previously used at Usuda for the tracking of the International Cometary Explorer (ICE) during the comet flyby period in 1985. This previous maser system installation has been thoroughly described in [3], so a detailed description will not be repeated here. Although the maser system was similar to the first implementation, modifications were made to improve its operation and reliability.

II. TWM/CCR

The TWM/CCR was originally built and implemented in the DSN in 1973 and was one of the first 2.3-GHz masers to

use the cold probe type of low-noise cryogenically cooled input transmission line [4]. Due to the age of the system and the amount of shipping and handling to which it had been exposed, it was felt necessary to open the CCR and make the desired corrections and improvements.

The signal input impedance match had degraded considerably in relation to its original performance. The input return loss was increased to 13 dB at 2270 MHz and 17.5 dB at 2300 MHz, a 3- to 4-dB improvement, by replacing and tuning a short section of coax in which there had been a slight migration of the Teflon dielectric. No changes were made to the maser structure itself.

The shock and vibration that had occurred during previous shipping had caused some hardware to become loose. Most notable were the screws that attach the thermal switch. A loose connection at this point causes a poor thermal contact and results in a longer maser cool-down period. The problem was attributed to rotational loads that could be applied to the joint during shipping, coupled with the soft copper material involved. An additional structural brace made of thin-wall

stainless steel tubing has been installed on two masers to date, and the problem has not recurred in the last 2 years.

New cryogenic temperature sensors were installed on the refrigerator, and the CCR controller in the instrumentation rack was modified to provide digital readout capability for the sensors. These sensors provide a continuous, accurate readout of refrigerator temperatures, enabling more accurate diagnosis of refrigerator performance problems than was previously possible. Identical sensors are now being installed in DSN maser systems.

III. Shipping, Installation, and Operation

The experience gained during the shipment of the TWM/CCR to Japan in 1985 helped prevent many of the problems that can occur due to the complexities of foreign customs policies. With this information available, it was possible to predict and eliminate all of the previous delays, and the system arrived in Usuda 8 days after its shipment from Los Angeles International Airport.

The beam-waveguide feed configuration of the Usuda antenna greatly contributed to the ease of installation of the TWM/CCR. Since most of the mounting hardware had been left in place in the laboratory environment of the beam-waveguide antenna, minimal problems or delays were encountered during installation. Whereas installation in a DSN Cassegrainian antenna requires advance scheduling of an 8-hour maintenance period, installation of the TWM/CCR system was completed and the cryogenics were ready to start one day after its arrival, with no interruption of ongoing tracking activities or maintenance.

The initial operation of the system was not without problems. One compressor and a CCR drive unit were replaced during the first week. After that, the maser continued to be operational without further mishaps during the 1-month-long TDRSS experiment 1986 observing schedule. In January 1987, during the second phase of the experiment, the TWM/CCR was easily restarted and operated without difficulty during the 10-day 1987 observing schedule.

The maser was tuned to 2277 MHz, and the initial system gain and power levels were adjusted. System temperature measurements were first made on July 3, 1986. The gain was adjusted to 50 dB with a bandwidth of approximately 20 MHz at the -3 dB points. Using the cold sky/ambient load Y-factor method, the system noise temperature was measured to be 14.9 K at an antenna elevation of 90 degrees. During January 1987, the maser gain was adjusted to 48 dB at a bandwidth of approximately 25 MHz. The system temperature was measured to be 15 K. This performance was essentially identical to data taken in 1985 for the ICE installation, when system noise temperature measured 15 K at 2270 and 2295 MHz.

IV. Conclusion

For the second time in as many years, a complete TWM/CCR system was shipped and installed on a foreign antenna in a timely manner. The project was a relatively cost-effective one, since it largely involved preexisting hardware. The TWM/CCR system performed reliably without the benefit of an operational spare and contributed to the success of the VLBI experiment it supported. The beam-waveguide feed configuration of the Usuda 64-m antenna, in contrast to the DSN Cassegrainian antennas, also contributed to ease of installation and operation.

Acknowledgments

The author wishes to express gratitude to the many people who made possible the success of this project. In particular, thanks are due to M. Britcliffe, whose knowledge and help in the timely diagnosis and correction of the initial cryogenic problems were indispensable, and to D. Neff, who was responsible for the previous installation and who furnished much help and guidance. Thanks also to L. Weight and his group for the packing and transportation between JPL and Usuda, and to R. Tanaka and his crew from Keihin-Burlington for their work in unpacking, repacking, and transporting the system within Japan. Special thanks to the many people from ISAS and UDSC for their help and cooperation in making the project a success.

References

- [1] G. Levy *et al.*, "Very Long Baseline Interferometric Observations Made With an Orbiting Radio Telescope," *Science*, vol. 234, pp. 187-189, October 10, 1986.
- [2] J. S. Ulvestad *et al.*, "Very Long Baseline Interferometry Using a Radio Telescope in Earth Orbit," *TDA Progress Report 42-88*, vol. October-December 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 1-10, February 15, 1987.
- [3] D. Neff, "Use of a 2.3-GHz Traveling-Wave Maser on the Usuda 64-Meter Antenna," *TDA Progress Report 42-89*, vol. January-March 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 34-40, May 15, 1987.
- [4] R. Clauss and E. Weibe, *Low-Noise Receivers: Microwave Maser Development*, JPL Technical Report 32-1526, vol. XIX, November-December 1973, Jet Propulsion Laboratory, Pasadena, California, pp. 93-99, February 15, 1974.

A Lower Bound for the Decoder Error Probability of the Linear MDS Code

K.-M. Cheung

Communications Systems Research Section

In this article, a lower bound for the decoder error probability ($P_E[u]$) of a linear maximum distance separable (MDS) code is derived by counting the dominant types of decoding words around code words. It is shown that the lower bound derived in this article is similar in form, and close numerically, to the upper bound derived in [2].

I. Introduction

Let C be a linear code of length n , dimension k , and minimum distance d . Let q be a positive power of a prime. An (n, k, d) linear code C over $GF(q)$ is *maximum distance separable* (MDS) if the Singleton bound is achieved; that is, $d = n - k + 1$. A code is t -error correcting if for some integer t , $2t \leq d - 1$.

In [1], by repeated use of the inclusion and exclusion principle, an exact expression for D_u , the number of decodable words of weight u , is derived. Also in [1], the exact decoding error probability $P_E(u)$ of a linear MDS code is evaluated. However, the formulas derived in [1] are complicated and clumsy, and offer no mathematical insight. In this article, by assuming that $q \geq n$, the lower bounds of $P_E(u)$ and $D(u)$ are derived from a completely different approach—simply by counting the dominant types of decoding words around code words. In Sections II and III the lower bound derived in this article is shown to be similar in form, and close numerically, to the upper bound derived in [2]. In Section IV, with the assumption that $q \geq n$, the lower bound of $P_E(u)$ as a func-

tion of u is shown to achieve its minimum value at $u = d - t$. Thus, the lower bound for $u = d - t$ is the overall lower bound of $P_E(u)$. For $q < n$, this may not be true.

II. Lower Bound of the Number of Code Words of Weight w

Let A_w denote the number of code words of weight w . A lower bound of A_w is given by the following theorem:

Theorem 1:

$$A_w \geq C \binom{n}{w} q^{-d+1} (q-1)^w \quad d \leq w \leq n \quad (1)$$

where

$$C = 1 - \frac{\binom{d}{2} q^{d-2}}{(q-1)^d}$$

Proof: From [1], A_w is given by the following expression:

$$\begin{aligned}
A_w &= \binom{n}{w} (q-1) \sum_{i=0}^{w-d} (-1)^i \binom{w-1}{i} q^{w-d-i} \\
&= \binom{n}{w} (q-1) q^{-d+1} \sum_{i=0}^{w-d} (-1)^i \binom{w-1}{i} q^{w-1-i} \\
&= \binom{n}{w} (q-1) q^{-d+1} \left[(q-1)^{w-1} \right. \\
&\quad \left. - \sum_{i=w-d+1}^{w-1} (-1)^i \binom{w-1}{i} q^{w-1-i} \right]
\end{aligned}$$

Consider the second term of the above expression. Since $q \geq n$,

$$\binom{w-1}{i} q^{w-1-i} \geq \binom{w-1}{i+1} q^{w-1-i-1}$$

for $d \leq w \leq n$ and $w-d+1 \leq i \leq w-1$. It is not hard to see that the following inequalities are obtained:

$$A_w \geq \binom{n}{w} q^{-d+1} (q-1)^w \quad (2)$$

$$A_w \leq \binom{n}{w} q^{-d+1} (q-1)^w \left[1 + \frac{\binom{w-1}{w-d+1} q^{d-2}}{(q-1)^{w-1}} \right]$$

$$w = d, d+2, d+4, \dots \quad (3)$$

and

$$A_w \leq \binom{n}{w} q^{-d+1} (q-1)^w \quad (4)$$

$$A_w \geq \binom{n}{w} q^{-d+1} (q-1)^w \left[1 - \frac{\binom{w-1}{w-d+1} q^{d-2}}{(q-1)^{w-1}} \right]$$

$$w = d+1, d+3, d+5, \dots \quad (5)$$

Consider the bracketed term in Eq. (5). Since $q \geq n$, it is an ascending function of w . So if we denote

$$\begin{aligned}
C &= \left[1 - \frac{\binom{w-1}{w-d+1} q^{d-2}}{(q-1)^{w-1}} \right]_{w=d+1} \\
&= 1 - \frac{\binom{d}{2} q^{d-2}}{(q-1)^d}
\end{aligned}$$

we have

$$A_w \geq C \binom{n}{w} q^{-d+1} (q-1)^w \quad d \leq w \leq n$$

where C is a scaling factor very close to 1. ■

III. Derivation of Lower Bound

Let \bar{d} be a decodable word. Then \bar{d} can be expressed uniquely as a sum $\bar{c} + \bar{e}$, where \bar{c} is a code word and \bar{e} is an error pattern of weight less than or equal to t . Let \bar{d} have weight u and \bar{e} have weight s , $s \leq t$. The weight of \bar{c} is then confined within a certain set of values, depending on the value of u and s . The main idea of deriving the lower bound of the number of decodable words of weight u is to count a certain "dominant" subset of code words that, when added to appropriate error patterns, gives rise to decodable words of weight u . Let us define

$$B_{u,s} = \{w : w \text{ is the weight of a code word that is at a distance } s \text{ from a decodable word of weight } u\}$$

We then have the following expression for $B_{u,s}$ depending on the value of u and s :

$$(1) \text{ If } d-t \leq u \leq d-1 \leq n-t, \\ \text{then } B_{u,s} = \{w : d \leq w \leq u+s\}$$

$$(2) \text{ If } d \leq u \leq d+t-1 \leq n-t, \\ \text{then } B_{u,s} = \{w : d \leq w \leq u+s\}$$

$$(3) \text{ If } d+t \leq u \leq n-t, \\ \text{then } B_{u,s} = \{w : u-s \leq w \leq u+s\}$$

$$(4) \text{ If } n-t+1 \leq u \leq n:$$

$$\text{If } u+s \leq n, \text{ then } B_{u,s} = \{w : u-s \leq w \leq u+s\}$$

$$\text{If } u+s > n, \text{ then } B_{u,s} = \{w : u-s \leq w \leq n\}$$

We can then express D_u as follows:

$$D_u = \sum_s \sum_{w \in B_{u,s}} A_w \times \{\# \text{ of error patterns of weight } s \text{ that give rise to a decodable word of weight } u \text{ from a code word of weight } w\} \quad (6)$$

We see that in the case $d - t \leq u \leq d - 1$, an allowable error pattern must be of weight $s \in \{d - u, \dots, t\} \subset \{0, 1, \dots, t\}$. In the case $d \leq u \leq n$, an allowable error pattern must be of weight $s \in \{0, 1, \dots, t\}$.

We also observe that for a linear MDS code, if $q \geq n$ and q is large, then

$$\frac{A_w}{A_{w-1}} \gg 1 \quad \text{for most } d \leq w \leq n$$

Thus, for the purpose of finding a lower bound of D_u , we do not need to consider all $w \in B_{u,s}$. We need only count those w 's that give rise to most decodable words of weight u . It is then logical to consider only those $w \in B'_{u,s} \subseteq B_{u,s}$ where $B'_{u,s}$ is a subset of $B_{u,s}$ ($B'_{u,s}$ consists of the larger numbers in $B_{u,s}$), instead of all $w \in B_{u,s}$. We now define $B'_{u,s}$ as follows:

- (1) If $d - t \leq u \leq d - 1 \leq n - t$,
then $B'_{u,s} = \{w: d \leq w \leq u + s\}$
- (2) If $d \leq u \leq d + t - 1 \leq n - t$,
then $B'_{u,s} = \{w: u \leq w \leq u + s\}$
- (3) If $d + t \leq u \leq n - t$
then $B'_{u,s} = \{w: u \leq w \leq u + s\}$
- (4) If $n - t + 1 \leq u \leq n$:
If $u + s \leq n$, then $B'_{u,s} = \{w: u \leq w \leq u + s\}$
If $u + s > n$, then $B'_{u,s} = \{w: u \leq w \leq n\}$

Before we proceed, we want to categorize the decodable words according to the following definition.

Definition 1:

Let \bar{d} be a decodable word which can be expressed in the form $\bar{d} = \bar{c} + \bar{e}$. Let $T_{\bar{c}}$ denote the set of nonzero coordinates of \bar{c} and $T_{\bar{e}}$ denote the set of nonzero coordinates of \bar{e} .

- (1) \bar{d} is defined to be of type A if $T_{\bar{e}} \subset T_{\bar{c}}$.
- (2) \bar{d} is defined to be of type B if it is not of type A.

It can be shown that for a given u , the number of type-A decodable words of weight u is usually much greater than the

number of type-B decodable words of weight u for most u . However, an explanation of the above claim is complicated and clumsy, and it is very hard to present a formal proof. A crude and oversimplified explanation is that type-A decodable words lie within Hamming spheres of code words of weights up to $u + t$, whereas type-B decodable words lie in the Hamming sphere of code words of weights only up to $u + t - 2$. As was mentioned before, $A_w \gg A_{w-1}$ for most w . This partly explains why the number of type-A decodable words is much greater than the number of type-B decodable words of weight u .

Summing up the above results, a lower bound of the number of decodable words of weight u is given by the following expression:

$$D_u \geq \sum_s \sum_{w \in B'_{u,s}} A_w \times \{\# \text{ of error patterns of weight } s \text{ that give rise to a type-A decodable word of weight } u \text{ from a code word of weight } w\} \quad (7)$$

We have four cases to consider, depending on the value of u .

- (1) $d - t \leq u \leq d - 1$

In this case, $s \in \{d - u, \dots, t\}$ and $w \in B'_{u,s} = \{d, d + 1, \dots, u + s\}$. There are $\binom{w}{s}$ ways of choosing s coordinates that give rise to type-A decodable words. But in order to have a type-A decodable word of weight u , the $w - u$ nonzero coordinates in \bar{e} must match with the corresponding $w - u$ nonzero coordinates in \bar{c} to give $w - u$ zeros in these coordinates. The remaining $s - (w - u)$ coordinates of \bar{e} must also match the corresponding $s - (w - u)$ coordinates of \bar{c} to give a nonzero value in each of the $s - (w - u)$ coordinates. There are $(q - 2)^{s - (w - u)}$ ways to do so.

Thus, the number of decodable words of weight u , where $d - t \leq u \leq d - 1$, is lower bounded as follows:

$$D_u \geq \sum_{s=d-u}^t \sum_{w \in B'_{u,s}} A_w \binom{w}{s} \binom{s}{w-u} (q-2)^{s-(w-u)}$$

We then substitute the lower bound of A_w in Eq. (1) for the above expression, and we have a lower bound of D_u as follows:

$$D_u \geq \sum_{s=d-u}^t \sum_{w=d}^{u+s} C \binom{n}{w} q^{-d+1} (q-1)^w \times \binom{w}{s} \binom{s}{w-u} (q-2)^{s-(w-u)}$$

We see that

$$\binom{n}{w} \binom{w}{s} \binom{s}{w-u}$$

can be expressed as

$$\binom{n}{u} \binom{n-u}{w-u} \binom{u}{s-(w-u)}$$

Let $\lambda = w - u$. The above expression can be rewritten as

$$D_u \geq \sum_{s=d-u}^t \sum_{\lambda=d-u}^s C \left(\frac{q-1}{q}\right)^{d-1} \left(\frac{q-2}{q-1}\right)^{s-w+u} \\ \times \binom{n}{u} \binom{n-u}{\lambda} \binom{u}{s-\lambda} (q-1)^{u+s-d+1}$$

Next, it is not hard to see that for the given ranges of u , s and w ,

$$\left(\frac{q-2}{q-1}\right)^{s-w+u} > \left(\frac{q-2}{q-1}\right)^t$$

Also, for the purpose of consistency with the equations that follow, the lower limit of the first summation on the RHS of the above expression can be replaced with 0 and thus our final expression is

$$D_u \geq C \left(\frac{q-1}{q}\right)^{d-1} \left(\frac{q-2}{q-1}\right)^t \binom{n}{u} (q-1)^{u-d+1} \\ \times \sum_{s=0}^t \sum_{\lambda=d-u}^s \binom{n-u}{\lambda} \binom{u}{s-\lambda} (q-1)^s$$

where

$$C = 1 - \frac{\binom{d}{2} q^{d-2}}{(q-1)^d}$$

$$(2) \quad d \leq u \leq d+t$$

In this case $s \in \{0, 1, \dots, t\}$ and $w \in \{u, u+1, \dots, u+s\}$. The derivation of lower bound of the number of decodable words of weight u is very similar to case 1, and the details of derivation are omitted. Since the smallest value of the

code word weights that are involved in counting is u , the scaling factor of the lower bound is now

$$C' = 1 - \frac{\binom{u-1}{u-d+1} q^{d-2}}{(q-1)^{u-1}}$$

which is closer to 1 than C . The lower bound of D_u is then given by

$$D_u \geq C' \left(\frac{q-1}{q}\right)^{d-1} \left(\frac{q-2}{q-1}\right)^t \binom{n}{u} (q-1)^{u-d+1} \\ \times \sum_{s=0}^t \sum_{\lambda=0}^s \binom{n-u}{\lambda} \binom{u}{s-\lambda} (q-1)^s$$

The lower bound can again be simplified by recalling the famous combinatoric identity

$$\sum_{\lambda=0}^s \binom{n-u}{\lambda} \binom{u}{s-\lambda} = \binom{n}{s}$$

and the final expression for this case is

$$D_u \geq C' \left(\frac{q-1}{q}\right)^{d-1} \left(\frac{q-2}{q-1}\right)^t \binom{n}{u} (q-1)^{u-d+1} \\ \times \sum_{s=0}^t \binom{n}{s} (q-1)^s \\ = C' \left(\frac{q-1}{q}\right)^{d-1} \left(\frac{q-2}{q-1}\right)^t \binom{n}{u} (q-1)^{u-d+1} \\ \times V_n(t) \quad d \leq u \leq d+t$$

$$(3) \quad d+t+1 \leq u \leq n-t$$

In this case, $s \in \{0, \dots, t\}$ and $w \in \{u, \dots, u+s\}$. The derivation is exactly the same as in case 2, and the lower bound is given by

$$D_u \geq C' \left(\frac{q-1}{q}\right)^{d-1} \left(\frac{q-2}{q-1}\right)^t \binom{n}{u} (q-1)^{u-d+1} \\ \times V_n(t) \quad d+t+1 \leq u \leq n-t$$

$$(4) \quad n-t+1 \leq u \leq n$$

In this case, if $u + s \leq n$ then $w \in \{u, \dots, u + s\}$, and if $u + s > n$ then $w \in \{u, \dots, n\}$. The derivation of the lower bound is slightly different from those of cases 2 and 3, but the final expression turns out to be the same. That is,

$$D_u \geq C' \left(\frac{q-1}{q}\right)^{d-1} \left(\frac{q-2}{q-1}\right)^t \binom{n}{u} (q-1)^{u-d+1} \\ \times V_n(t) \quad n-t+1 \leq u \leq n$$

In summary, the lower bound of the number of decodable words is given by the following equations:

$$D_u \geq C \left(\frac{q-1}{q}\right)^{d-1} \left(\frac{q-2}{q-1}\right)^t \binom{n}{u} (q-1)^{u-d+1} \\ \times \sum_{s=0}^t \sum_{\lambda=d-u}^s \binom{n-u}{\lambda} \binom{u}{s-\lambda} (q-1)^s \\ d-t \leq u \leq d-1 \quad (8)$$

$$D_u \geq C' \left(\frac{q-1}{q}\right)^{d-1} \left(\frac{q-2}{q-1}\right)^t \binom{n}{u} (q-1)^{u-d+1} \\ \times V_n(t) \quad n-t+1 \leq u \leq n \quad (9)$$

where

$$C = 1 - \frac{\binom{d}{2} q^{d-2}}{(q-1)^d}$$

and

$$C' = 1 - \frac{\binom{u-1}{u-d+1} q^{d-2}}{(q-1)^{u-1}}$$

We have shown in [1] that the decoder error probability is related to the number of decodable words via Eq. (2) and thus the decoder error probability $P_E(u)$ is lower bounded as follows:

$$P_E(u) \geq C q^{-d+1} \left(\frac{q-2}{q-1}\right)^t \sum_{s=0}^t \sum_{\lambda=d-u}^s \binom{n-u}{\lambda} \binom{u}{s-\lambda} (q-1)^s \\ d-t \leq u \leq d-1 \quad (10)$$

$$P_E(u) \geq C' q^{-d+1} \left(\frac{q-2}{q-1}\right)^t V_n(t) \quad d \leq u \leq n \quad (11)$$

where

$$C = 1 - \frac{\binom{d}{2} q^{d-2}}{(q-1)^d}$$

and

$$C' = 1 - \frac{\binom{u-1}{u-d+1} q^{d-2}}{(q-1)^{u-1}}$$

IV. Overall Lower Bound of $P_E(u)$

In this section, an overall lower bound of $P_E(u)$ for all u is given by the following theorem and corollary.

Theorem 2:

If $q \geq n$, then the lower bound of $P_E(u)$ in Eqs. (10) and (11) is smallest for $u = d - t$.

Proof: First of all, it is not hard to see that the lower bound in Eq. (10) is always smaller than the lower bound in Eq. (11) because $\binom{n}{s}$ is always greater than the incomplete Vandermonde convolution $\sum_{\lambda=d-u}^t \binom{n-u}{\lambda} \binom{u}{t-\lambda}$. Also, the scaling factor C' in Eq. (11) is always greater than the scaling factor C in Eq. (10). Thus, to prove the theorem, we need only consider the lower bound of $P_E(u)$ for $d-t \leq u \leq d-1$. It is not hard to see that a sufficient condition is to show that

$$\sum_{s=0}^t \sum_{\lambda=d-u}^s \binom{n-u}{\lambda} \binom{u}{s-\lambda} (q-1)^s \geq \binom{n-d+t}{t} (q-1)^t$$

$$d-t \leq u \leq d-1$$

It is obvious that

$$\sum_{s=0}^t \sum_{\lambda=d-u}^s \binom{n-u}{\lambda} \binom{u}{s-\lambda} (q-1)^s \\ \geq \sum_{\lambda=d-u}^t \binom{n-u}{\lambda} \binom{u}{t-\lambda} (q-1)^t$$

We now proceed to show that

$$\sum_{\lambda=d-u}^t \binom{n-u}{\lambda} \binom{u}{t-\lambda} (q-1)^t \geq \binom{n-d+t}{t} (q-1)^t$$

Let $l = t - d + u$ and $m = t - \lambda$; we have

$$\begin{aligned} \sum_{\lambda=d-u}^t \binom{n-u}{\lambda} \binom{u}{t-\lambda} (q-1)^t &= \sum_{m=0}^l \binom{n-d+t-l}{t-m} \\ &\times \binom{d-t+l}{m} (q-1)^t \end{aligned}$$

Since $d \geq 2t + 1$ and $0 \leq l \leq t - 1$,

$$\binom{d-t+l}{m} \geq \binom{l}{m}$$

Thus,

$$\begin{aligned} &\sum_{m=0}^l \binom{n-d+t-l}{t-m} \binom{d-t+l}{m} (q-1)^t \\ &\geq \sum_{m=0}^l \binom{n-d+t-l}{t-m} \binom{l}{m} (q-1)^t = \binom{n-d+t}{t} (q-1)^t \end{aligned}$$

and the theorem is proved. ■

Corollary :

An overall lower bound of $P_E(u)$ for all u is

$$\begin{aligned} P_E(u) &\leq C \left(\frac{q-2}{q-1}\right)^t \left(\frac{q-1}{q}\right)^{d-1} P_E(d-t) \\ &= C \left(\frac{q-2}{q-1}\right)^t q^{-d+1} \binom{n-d+t}{t} (q-1)^t \end{aligned}$$

where

$$C = 1 - \frac{\binom{d}{2} q^{d-2}}{(q-1)^d}$$

Proof: A direct result from Theorem 2. ■

V. Remarks

For $q \geq n$, the upper bound and lower bound of $P_E(u)$ give a good estimation of $P_E(u)$. The upper bounds [2], lower bounds, and exact values of the $P_E(u)$'s of the NASA code and the JTIDS code are tabulated in Table 1 and Table 2, respectively. We observe that the estimated values (upper bound and lower bound) are more or less of the same order of magnitude as the exact value in each case.

Also, we have shown that with the assumption that $q \geq n$, an overall lower bound of $P_E(u)$ (for all u) is given by

$$C \left(\frac{q-2}{q-1}\right)^t P_E(d-t)$$

■ For $q < n$, this may not be true.

References

- [1] K.-M. Cheung, "More on the Decoder Error Probability for Reed-Solomon Codes," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 213-221, November 15, 1987.
- [2] R. J. McEliece and L. Swanson, "On the Decoder Error Probability for Reed-Solomon Codes," *IEEE Tran. Inform. Theory*, vol. IT-32, pp. 701-703, 1986.

Table 1. Decoder error probability of the NASA code*

Weight	Lower bound	Actual value	Upper bound
17	7.769×10^{-15}	9.464×10^{-15}	2.956×10^{-14}
18	1.665×10^{-14}	1.913×10^{-14}	2.957×10^{-14}
19	2.171×10^{-14}	2.401×10^{-14}	2.957×10^{-14}
20	2.361×10^{-14}	2.660×10^{-14}	2.957×10^{-14}
21	2.414×10^{-14}	2.602×10^{-14}	2.957×10^{-14}
22	2.425×10^{-14}	2.608×10^{-14}	2.957×10^{-14}
.	.	.	.
.	.	.	.
.	.	.	.
37	2.450×10^{-14}	2.609×10^{-14}	2.957×10^{-14}
.	.	.	.
.	.	.	.

*NASA code (255, 223); $q = 256$; $t = 16$.

Table 2. Decoder error probability of the JTIDS code*

Weight	Lower bound	Actual value	Upper bound
9	1.340×10^{-6}	3.750×10^{-6}	9.250×10^{-6}
10	5.741×10^{-6}	1.439×10^{-6}	9.349×10^{-6}
11	1.310×10^{-6}	2.951×10^{-6}	9.350×10^{-6}
12	2.123×10^{-6}	4.329×10^{-6}	9.350×10^{-6}
13	2.767×10^{-6}	5.189×10^{-6}	9.350×10^{-6}
14	3.140×10^{-6}	5.547×10^{-6}	9.350×10^{-6}
.	.	.	.
.	.	.	.
.	.	.	.
25	4.328×10^{-6}	5.626×10^{-6}	9.350×10^{-6}
.	.	.	.
.	.	.	.

*RS code (31, 15); $q = 32$; $t = 8$.

A Labeling Procedure for Linear Finite-State Codes

K.-M. Cheung

Communications Systems Research Section

In this article a method to define the labels of the state diagram of a linear finite-state code [1] is presented and investigated. This method is particularly suitable for simple hardware implementation since it simplifies the encoder structure. The method can also be applied to the labeling of a state diagram that is not completely connected to obtain a linear finite state code with larger free distance.

I. Introduction

It was shown in [1] that a finite-state code (*FS* code) on a completely connected state diagram with 2^m states requires at least 2^{m+1} labels. Also, a simple method to define such labels has been suggested in [1]. However, the codes constructed using the method in [1] are not linear. In this article, another method using shift registers to define the labels of the state diagram of the *FS* codes is presented. This method is particularly suitable for simple hardware implementations since it simplifies the encoder structure. The method can also be applied to the labeling of a state diagram that is not completely connected to obtain an *FS* code with larger free distance. Lastly, a mapping scheme to assign the cosets to the labels generated by the shift registers is described. It can be shown that by using the above method, a linear *FS* code can be constructed.

In order to facilitate the discussion on *FS* codes with non-completely connected state diagrams as well as those with completely connected state diagrams, the following definition of *FS* codes is adopted:

Definition 1: An (n, k, m) finite state code (*FS* code) on a c -connected state diagram is a code with the following properties:

- (1) The code has rate k/n .
- (2) Its operation can be represented by a state diagram with 2^m states.
- (3) There are 2^c ($c \leq m$) branches going into each state and 2^c branches going out of each state.
- (4) Each branch of the state diagram is associated with a code (code word length = n and code size = 2^{k-c}), and any two different codes associated with different branches are disjoint.

II. Preliminaries

Some important results in the theory of convolutional codes will now be reviewed. These results will be referred to in the proofs in later sections.

A typical encoder of an (n_1, c, m) convolutional code consists of a linear sequential circuit (with c shift registers) that accepts c input bits and outputs n_1 bits. It is well known that the operation of the encoder can be represented by (1) a state diagram with 2^m states, 2^c branches going into each state, and 2^c branches going out of each state; or (2) a $c \times n_1$ transfer function matrix (denoted by $G[D]$) such that the entries of the matrix are polynomials in D , representing the generator sequences of the code.

In order to avoid catastrophic error propagation, the transfer function matrix must satisfy Massey and Sain's condition [2] (a necessary and sufficient condition) on non-catastrophic codes:

$$GCD \left[\Delta_i(D), \quad i = 1, 2, \dots, \binom{n_1}{c} \right] = D^l$$

for some $l \geq 0$, where $\Delta_i(D)$, $i = 1, 2, \dots, \binom{n_1}{c}$ are the determinants of the $\binom{n_1}{c}$ distinct $c \times c$ submatrices of the transfer function matrix $G(D)$.

III. Generation of Labels by Shift Register

FS code encoders have structural properties very similar to those of convolutional encoders, and their operation can be described by a state diagram. In the case of a convolutional code, each branch of the state diagram is labeled by an n_1 -bit output sequence, whereas in the case of a finite-state code according to Definition 1, each branch is labeled by a code that is not necessarily linear. Because of the similarities between convolutional codes and finite state codes, it should be expected that much of the theory on structural properties of convolutional codes will be applicable to finite state codes.

In order to guarantee a noncatastrophic finite state code with good distance properties, the labeling of the branches of the state diagram must satisfy the following conditions [1]: (1) different labels out of each state; (2) different labels into each state; and (3) no disjoint paths with identical labels that remain unmerged indefinitely.

A method to assign the labels of the state diagram of a finite state code by using the linear sequential circuit (with shift registers) of a noncatastrophic (n_1, c, m) convolutional code is now described. Let the c shift registers have lengths l_1, l_2, \dots, l_c where $l_1 + l_2 + \dots + l_c = m$. The p th row of the corresponding $c \times n_1$ transfer function matrix thus consists of polynomials in D of degree no greater than l_p for $1 \leq p \leq c$. The state diagram of the convolutional code consists of 2^m states (each state is defined by the shift register content);

also, there are 2^c branches going into each state and 2^c branches going out of each state. Each branch in the state diagram is assigned an n_1 -bit sequence $b_0, b_1, \dots, b_{n_1-1}$, which consists of the n_1 output bits of the shift registers. Let us assign to the branches of the state diagram, which are associated with the n_1 -bit sequence $b_0, b_1, \dots, b_{n_1-1}$, the label i such that $i = b_0 + 2b_1 + \dots + 2^{n_1-1} b_{n_1-1}$. Each of these labels represents one of the disjoint codes. There are 2^{n_1} of them. This modified state diagram of the convolutional code is used as the state diagram of an (n, k, m) finite state code on a c -connected state diagram.

The construction of a shift register circuit that generates the state diagram of a finite state code that satisfies conditions 1, 2, and 3 is given as follows. It is not hard to see that condition 1 is satisfied if, for a fixed shift register content, different inputs to the shift registers produce different outputs. This can be achieved if there exists at least one $c \times c$ submatrix $\Omega_i(D)$ of the transfer function matrix $G(D)$, $i = 1, 2, \dots, \binom{n_1}{c}$, such that the term "1" appears exactly once in each row and in each column of $\Omega_i(D)$. Similarly, condition 2 is satisfied if, for a fixed input, different shift register contents produce different outputs. This can be achieved if there exists at least one $c \times c$ submatrix $\Omega_j(D)$, $j = 1, 2, \dots, \binom{n_1}{c}$, such that the term D^{lp} representing the last shift register stage of the p th shift register appears exactly once in row p for $1 \leq p \leq c$, and each of these $D^{l_1}, D^{l_2}, \dots, D^{l_c}$ terms appears in different columns of $\Omega_j(D)$.

It was shown in [3] that if the (n_1, c, m) convolutional code that generates the state diagram of the finite state code is noncatastrophic, then the labeling also satisfies condition 3. Thus, the $c \times n_1$ transfer function matrix $G(D)$ of the convolutional code must satisfy Massey and Sain's condition [2]. It will be shown in later sections that the minimum value n_1 could have is $c + 1$. Two algorithms to construct a $c \times c + 1$ transfer function matrix $G(D)$ of the convolutional code are given as follows:

Algorithm 1: Completely connected state diagram, $d_f = 2$ branches.

(1) Construct a $c \times c$ matrix $G'(D)$ such that

$$\begin{aligned} G'_{ij}(D) &= 1, & i &= 1, \dots, c, & j &= i \\ &= D, & i &= 1, \dots, c, & j &= (i+1) \bmod c \\ &= 0 & & \text{otherwise} \end{aligned}$$

(2) Append the column $[1, \dots, 0]^T$ to $G'(D)$ to obtain a $c \times c + 1$ matrix $G(D)$.

An example of a 3×4 transfer function matrix $G(D)$ constructed using the above algorithm is given in Table 1. It is obvious that $G(D)$ satisfies conditions 1 and 2. Also, it is not hard to see that one of the determinants, $\Delta_j(D)$, $j = 1, \dots, (c+1)$, equals 1 and the rest are nonzero. Thus

$$\text{GCD} \left[\Delta_j(D), \quad j = 1, \dots, \binom{c+1}{c} \right] = 1$$

Massey and Sain's condition is satisfied and the state diagram generated by this transfer function matrix satisfies conditions 1, 2, and 3.

Algorithm 2: Non-completely connected state diagram, $d_f = 3$ branches.

(1) Construct a $c \times c$ matrix $G'(D)$ such that

$$\begin{aligned} G'_{ij}(D) &= 1 + D, & i = 1, \dots, c, & \quad j = i \\ &= D^2, & i = 1, \dots, c, & \quad j = (i+1) \bmod c \\ &= 0 & \text{otherwise} \end{aligned}$$

(2) Append the column $[1, \dots, 0]^T$ to $G'(D)$ to obtain a $c \times c + 1$ matrix $G(D)$.

An example of a 3×4 transfer function matrix $G(D)$ constructed using the above algorithm is given in Table 2. Again it is obvious that $G(D)$ satisfies conditions 1 and 2. Also, it can be shown that one of the determinants, $\Delta_j(D)$, equals $D^{2(c-1)}$ and the rest are nonzero. Thus

$$\text{GCD} \left[\Delta_j(D), \quad j = 1, \dots, \binom{c+1}{c} \right] = D^l$$

where l is some integer. Massey and Sain's condition is satisfied and the state diagram generated by $G(D)$ satisfies conditions 1, 2, and 3.

IV. Properties

On the basis of the labeling procedure by shift register above, which is based on a linear sequential circuit, the finite state code possesses a mathematical structure that facilitates encoding/decoding and simplifies hardware implementation. Also, this labeling procedure is applicable to the construction of finite state codes with incompletely connected state diagrams to obtain larger free distance.

Definition 2: Let N be the number of states of a finite state code. A labeling matrix L of the state diagram is defined to be

an $N \times N$ matrix, where $L(i, j)$ denotes the label from state i to state j .

Let $\underline{u} = (u_1, u_2, \dots, u_c)$ represent the c input bits to the convolutional encoder. Let $\underline{D} = (D_1, D_2, \dots, D_c)$ represent the last c shift register stages of the convolutional encoder. That is, D_p represents the term D^p in row p for $1 \leq p \leq c$. In the following theorems, some properties of FS codes which use the new labeling procedure are revealed.

Theorem 1: For a state diagram with 2^m states generated by $G(D)$ which satisfies conditions 1, 2, and 3, if the graph has 2^c branches going into each state and 2^c branches going out of each state, $c \leq m$, at least 2^{c+1} labels are required.

Proof: Suppose that 2^c labels suffice. The transfer function matrix $G(D)$ of the convolution code that generates the state diagram of the FS code is then a $c \times c$ matrix. By condition 1, since different labels are coming out of each state, the c output bits can be written as

$$\underline{u}\mathbf{A} + \underline{d}$$

where \mathbf{A} is a $c \times c$ nonsingular matrix and \underline{d} is a constant binary c -tuple which depends upon the shift register contents of the encoder. Thus, $|\mathbf{A}| = \alpha$, where α is a nonzero integer. Thus, the term α is contained in the expression of $|\mathbf{M}|$. Similarly, by condition 2, since different labels are going into each state, the c output bits can be written as

$$\underline{D}\mathbf{B} + \underline{e}$$

where \mathbf{B} is a $c \times c$ nonsingular matrix and \underline{e} is a constant binary c -tuple, depending upon the input bits and the shift register contents other than D_1, \dots, D_c . Again, $|\mathbf{B}| = \beta$ for some nonzero integer β . Therefore the term $\beta D^m = \beta D^{l_1 + \dots + l_c}$ is contained in the expression of $|\mathbf{M}|$. Thus, $|\mathbf{M}| = \beta D^m + \dots + \alpha$ and $|\mathbf{M}|$ is not of the form D^l for some $l \geq 0$. This violates Massey and Sain's condition and the convolutional code is catastrophic. This in turn implies that the state diagram generated by this convolutional encoder is catastrophic and thus at least $c + 1$ output bits for the convolutional encoder are needed. This implies that at least 2^{c+1} labels are needed in the state diagram. ■

In fact, Algorithm 1 and Algorithm 2 in Section II show that $c + 1$ output bits are sufficient to guarantee that conditions 1, 2, and 3 are satisfied.

Theorem 2: Let L be the labeling matrix of a state diagram generated by $G(D)$ which satisfies conditions 1, 2, and 3. Row i and row j (column i and column j), $i \neq j$, of L have

either the same set of labels or a completely different set of labels.

Proof: The state of the convolutional encoder that generates the required state diagram of the finite state code is defined as the shift register contents of the encoder. For an (n_1, c, m) convolutional code, let the binary m -tuple $[D_1, \dots, D_m]$ denote the state that corresponds to the shift register stages D_1, \dots, D_m of the encoder. Note that the encoder is constructed in such a way that for a fixed state $[D_1, \dots, D_m]$, different inputs to the shift registers produce different outputs (condition 1). If $[D_1, \dots, D_m] = [0, \dots, 0]$, the set of all possible binary n_1 -tuples (labels) that represent the output bits of the encoder forms a c -dimension subspace K of an n_1 -dimension vector space over $GF(2)$ (because the encoder is a linear sequential circuit). This set K is isomorphic to the row of the labeling matrix L that corresponds to the state $[0, \dots, 0]$. Now, if $[D_1, \dots, D_m] \neq [0, \dots, 0]$, then it is not hard to see that the set of all possible output binary n_1 -tuples (output bits of the encoder) is of the form $K + \underline{e}$, where \underline{e} is a binary n_1 -tuple (constant) determined by $[D_1, \dots, D_m]$. If $\underline{e} \notin K$, then K and $K + \underline{e}$ are disjoint (since K is a c -dimensional subspace in an n_1 -dimensional vector space). If $\underline{e} \in K$, then $K = K + \underline{e}$. A similar argument holds for the case of $K + \underline{e}_1$ and $K + \underline{e}_2$, where \underline{e}_1 and \underline{e}_2 are binary n_1 -tuples determined by different $[D_1, \dots, D_m]$'s. That is, if $\underline{e}_1 \notin K + \underline{e}_2$, then $K + \underline{e}_1$ and $K + \underline{e}_2$ are disjoint. If $\underline{e}_1 \in K + \underline{e}_2$ then $K + \underline{e}_1 = K + \underline{e}_2$. This proves that any two rows of a labeling matrix L have either the same set of labels or a completely different set of labels. The proof for the case of the columns is similar to the one above. ■

V. Assignment of Cosets to Labels

A code C over $GF(q)$ is said to be linear if and only if the following condition is satisfied:

$$\forall \underline{a}, \quad \underline{b} \in C \text{ and } \forall \gamma, \quad \delta \in GF(q), \quad \gamma \underline{a} + \delta \underline{b} \in C$$

In an FS code, even though we have a linear convolutional structure (labels are generated by outputs of shift registers), the overall code may not be linear if the cosets are not properly assigned to the outputs of shift registers. There may exist two code word sequences such that their sum is not a legal code word sequence. In order to generate a linear FS code the following well-known theorem in linear algebra can be used:

Theorem 3 (without proof): If C is a vector space and S is a proper subspace of C , then there exists a subspace W of C such that

$$S + W = C$$

$$S \cap W = \{0\}$$

$$\dim S + \dim W = \dim C$$

The following discussion describes a way to generate a linear FS code. The labeling of an FS code can be divided into two parts: (1) generation of labels to the branches in the state diagram; and (2) assignment of cosets to the labels. Part 1 was taken care of by using a convolutional encoder to generate labels to the state diagram of the FS code. For part 2, the method proceeds as follows. Let C be the parent (n, k) code. Let S be an (n, k_1) subcode of C . By Theorem 3, there exists a subcode W of C (W is an $[n, k - k_1]$ code) such that

$$S + W = C$$

$$S \cap W = \{0\}$$

$$\dim S + \dim W = \dim C$$

The 2^{k-k_1} cosets are constructed by adding each word in W to S . That is,

$$\underline{w} + S \quad \forall \underline{w} \in W$$

Note that the set of all binary $k - k_1$ -tuples is isomorphic to W . Let $\{\underline{w}_0, \underline{w}_1, \dots, \underline{w}_{k-k_1-1}\}$ be a basis of W . Let $b_0, b_1, \dots, b_{k-k_1-1}$ be the $k - k_1$ output bits of the convolutional encoder. Let the coset assigned to the branches labeled by the binary $(k - k_1)$ -tuples $b_0, b_1, \dots, b_{k-k_1-1}$ be denoted by $L(b_0, b_1, \dots, b_{k-k_1-1})$. Let us assign

$$L(b_0, b_1, \dots, b_{k-k_1-1}) = S + \{b_0 \underline{w}_0 + b_1 \underline{w}_1 + \dots + b_{k-k_1-1} \underline{w}_{k-k_1-1}\}$$

This assignment of cosets to the branches in the state diagram guarantees the linearity of the FS code.

References

- [1] F. Pollara, R. McEliece, and K. Abdel-Ghaffar, "Constructions for Finite State Codes," *TDA Progress Report 42-90*, vol. April-June 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 42-49, August 15, 1987.
- [2] J. Massey and M. Sain, "Inverse of Linear Sequential Circuit," *IEEE Trans. Comput.*, vol. C-17, 1968.
- [3] K.-M. Cheung, "Error-Correction Coding for Data Storage Systems," PhD thesis, California Institute of Technology, 1987.

58-61
128224

N88-18782

TDA Progress Report 42-92

October-December 1987

Further Results on Finite-State Codes

F. Pollara and K.-M. Cheung
Communications Systems Research Section

R. J. McEliece
California Institute of Technology

A general construction for finite-state (FS) codes is applied to some well-known block codes. New subcodes of the (24,12) Golay code are used to generate two optimal FS codes with $d_{free} = 12$ and 16. A partition of the (16,8) Nordstrom-Robinson code yields a $d_{free} = 10$ FS code. Simulation results are shown and decoding algorithms are briefly discussed.

I. Introduction

Future deep-space communication systems will take advantage of powerful error-correcting coding schemes to keep power and antenna size requirements within acceptable bounds. Such codes can be found by computer search or, as considered in this article, by constructions based on known codes.

In a previous article [1] it was shown how some optimal Finite-State (FS) codes can be constructed from known block codes. This article considers new FS codes based on other block codes and describes performance results obtained by simulation.

II. Codes Derived From the (24,12) Golay Code

The basic idea developed in [1] consists in choosing an (n, k_1) block code C_1 with minimum distance d_1 , and then decomposing C_1 into the disjoint union of cosets generated by

an (n, k_2) subcode C_2 of C_1 , with minimum distance d_2 . By properly assigning these cosets to the edges of a 2^m -state completely connected graph, it is possible to construct an (n, k, m) FS code, with $k = m + k_2$ and $d_{free} \geq \min(d_2, 2d_1)$.

The (24,12) Golay code could be an interesting candidate for this construction provided that it contains a subcode with minimum distance d_2 larger than $d_1 = 8$. The following theorem shows that such a subcode does indeed exist.

Theorem 1. The (24,12) Golay code has a (24,5) subcode with minimum distance 12.

Proof: The proof is based on the Turyn construction of the Golay code (p. 587 of [2]). Let A be the (7,3) code with code words consisting of (0,0,0,0,0,0,0) and the seven cyclic shifts of (1,1,0,1,0,0,0). Then the (7,4) code $H = A \cup \bar{A}$, where the bar denotes the complemented code words, is the (7,4) Hamming code. Similarly, consider the code A^* obtained by reversing the order of symbols in A , and the code $H^* = A^* \cup \bar{A}^*$. Let

Table 1. Transfer function matrix to generate a completely connected state diagram with 8 states and 16 labels

	1	0	D	1	
$G(D) =$	D	1	0	0	$d_f = 2$ branches
	0	D	1	0	

Table 2. Transfer function matrix to generate a non-completely connected state diagram with 64 states and 16 labels

	$1+D$	0	D^2	1	
$G(D) =$	D^2	$1+D$	0	0	$d_f = 3$ branches
	0	D^2	$1+D$	0	

C and C^* be (8,4) codes obtained by adding a parity check bit to H and H^* . Then C and C^* have $d_{\min} = 4$, and the code G consisting of all vectors

$$|a+x|b+x|a+b+x|, \quad a, b \in C, \quad x \in C^* \quad (1)$$

is the (24,12) binary Golay code with $d_{\min} = 8$.

Let B be the subcode of C consisting of the two code words (0,0,0,0,0,0,0) and (1,1,1,1,1,1,1). Then the construction in (1) with $a, b \in B$ and $x \in A^*$ generates code words of the form

$$|x|x|x|, \quad |x|\bar{x}|\bar{x}|, \quad |\bar{x}|x|\bar{x}|, \quad |\bar{x}|\bar{x}|x| \quad (2)$$

Code words taken from two distinct subcodes of the four subcodes above are at minimum distance $8 + 8 = 16$ for fixed x , and at distance $4 \times 3 = 12$ for $x \neq y \in A^*$. Code words in the same subcode are at minimum distance $4 \times 3 = 12$. Therefore, by using all 8 possible choices for x , we have constructed a (24,5) subcode of the Golay code with $d_{\min} = 12$. ■

Previously known (24,5) subcodes of the Golay code have $d_{\min} = 8$ [3], [4]. The (24,5) subcode just described can be represented on a trellis as shown in Fig. 1, where each edge x or \bar{x} corresponds to eight bits. Figure 1 consists of the union of 8 cosets D_i , $i = 0, 1, \dots, 7$, given by (2) with $x \in A^*$. Each coset has 4 code words and is represented by a trellis as shown in Fig. 2. This observation leads to the following result.

Corollary 1. The (24,12) Golay code has a (24,2) subcode with minimum distance equal to 16.

Proof: This follows directly from Expression (2) with $x = (0,0,0,0,0,0,0)$. Figure 2 shows the trellis representing the (24,2) subcode with 4 code words. ■

Since there are 128 (24,5) cosets in the Golay code, it is possible to construct a non-catastrophic [1] FS code with up to 64 states on a completely connected graph. By this construction we obtain a (24,11,6) FS code with $d_{\text{free}} = \min(12, 2 \times 8) = 12$. For this code, since d_2 is strictly smaller than $2d_1$, it is also possible to say that there are exactly 30 error events at distance 12, the number of code words of the (24,5) subcode of weight 12. Similarly, by using the 2^{10} (24,2) cosets, we can construct a (24,11,9) FS code with $d_{\text{free}} = \min(16, 2 \times 8) = 16$. These new codes are both optimal in the sense that they achieve the largest possible free distance, as predicted by the Plotkin bound for FS codes [1].

III. Codes Derived From the Nordstrom–Robinson Code

In [1] a (16,7,2) FS code was constructed starting from the nonlinear (16,8) Nordstrom–Robinson code with $d_{\min} = 6$, which is the union of 8 particular cosets of the (16,5) first-order Reed–Müller code with $d_{\min} = 8$.

Given that the Nordstrom–Robinson code has many pairs of code words at distance 10 and that a (16, k) code may have $d_{\min} = 10$ only if $k \leq 2$ (by the Plotkin bound), it is interesting to see if the Nordstrom–Robinson code can be split into 64 sets of 4 code words, each with $d_{\min} = 10$. The following theorem proves that this is true.

Theorem 2. The (16,8) Nordstrom–Robinson code can be partitioned into 64 sets, each having $d_{\min} = 10$.

Proof: The Nordstrom–Robinson code is the union of 8 cosets of the (16,5) first-order Reed–Müller code. Let the 8 coset leaders be denoted by a_i and b_i , $i = 0, 1, 2, 3$. Then $a_0 = 0$ and the other coset leaders can be taken to be the following seven bent¹ functions of four Boolean variables x_1, x_2, x_3, x_4 (problem 21, p. 476 of [2]),

$$a_1 = x_1x_2 + x_1x_3 + x_2x_3 + x_2x_4$$

$$a_2 = x_1x_2 + x_3x_4$$

$$a_3 = x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$$

$$b_0 = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_4$$

$$b_1 = x_1x_2 + x_1x_4 + x_2x_3$$

$$b_2 = x_1x_3 + x_2x_4 + x_3x_4$$

$$b_3 = x_1x_3 + x_1x_4 + x_2x_3 + x_3x_4$$

If instead the coset leaders are taken to be:

$$A_0 = 0$$

$$A_1 = a_1 + x_3 + 1$$

$$A_2 = a_2 + x_2 + x_4 + 1$$

$$A_3 = a_3 + x_1 + x_2 + 1$$

$$B_0 = b_0 + x_2 + x_3 + x_4$$

$$B_1 = b_1 + x_1 + x_4$$

$$B_2 = b_2 + x_3 + x_4 + 1$$

$$B_3 = b_3 + x_1 + x_2 + x_3 + 1$$

¹These Boolean functions are so called because they are in some sense furthest away from linear functions.

then $\text{dist}(A_i, A_j) = 10$, $\text{dist}(B_i, B_j) = 10$, and $\text{dist}(A_i, B_j) = 6$ for all i and j , $i \neq j$. Then, for each word w in the (16,5) first-order Reed-Müller code ($A_0 + w, A_1 + w, A_2 + w, A_3 + w$) and ($B_0 + w, B_1 + w, B_2 + w, B_3 + w$) are subsets of the Nordstrom-Robinson code with distance 10. There are 64 such subsets and they exhaust the Nordstrom-Robinson code. ■

Table 1 shows the code words of the (16,5) first-order Reed-Müller code and the 8 coset leaders that generate the 64 subsets used for the FS code construction. By assigning the 64 subsets to the edges of a 32-state completely connected graph a (16,7,5) FS code can be constructed. This code has $d_{\text{free}} = \min(2 \times 6, 10) = 10$, which meets the Plotkin bound.

IV. Simulation Results and Decoding Algorithms

An existing software simulation for FS codes has been adapted to the newly found codes. Simulation results showing the probability of bit error versus E_b/N_0 are given in Fig. 3. The (24,11,6) FS code with $d_{\text{free}} = 12$ and the (16,7,5) FS code with $d_{\text{free}} = 10$ are compared for reference to the (2,1,6) Voyager convolutional code.

These results are obtained by a soft, maximum-likelihood decoder based on the Viterbi algorithm. The decoder performs two basic steps:

- (1) Each received word (24 or 16 symbols) is compared to the code words in each coset (128 or 64) and the closest code word in each coset is stored together with its distance.
- (2) At each state, the decoder further selects the closest code word among those chosen in step 1 for the cosets assigned to branches reaching that state.

For the (24,11,6) code, the total number of bit operations per decoded bit involved in the decoding process is $(24/11) 2^{12}$, where 2^{12} is the total number of branches in one stage of the decoder trellis. It is interesting to note that the same number for the Golay code is $(24/12) 2^{12}$, which is very close, but the FS code has $d_{\text{free}} = 12$ compared to a $d_{\text{free}} = 8$ of the Golay code. Similarly, the decoding of the (16,7,5) code involves $(16/7) 2^{10}$ bits per decoded bit.

V. Conclusion

In this article we have described FS codes based on partitions of the Golay and Nordstrom-Robinson codes, which did not appear in the literature.

The comparison of these new codes to known codes, block and convolutional, is complicated by the fact that both the performance and the decoding complexity must be taken into account, and the complexity is intimately related to the particular hardware architecture used for the decoder. We feel that the proposed codes may take greater advantage of parallel VLSI architectures than conventional convolutional codes with no structure. Also, the trellis representation of cosets as in Figs. 1 and 2 can be used to reduce the number of comparisons to select the closest code word with methods similar to those described in [3].

Figure 4 summarizes the present knowledge on FS codes by showing the Plotkin or Hamming bound (whichever is tighter) on the free distance achievable for a given encoder memory and for two classes of FS codes, the (24,11, m) and the (16,7, m) classes. The (16,7,2) code has been reported in [1]. The Voyager code is also shown for comparison as a member of the (2,1, m) class of convolutional codes. More work needs to be done in constructing yet more powerful FS codes, especially those based on graphs that are not completely connected.

References

- [1] F. Pollara, R. J. McEliece, and K. Abdel-Ghaffar, "Constructions for Finite-State Codes," *TDA Progress Report 42-90*, April-June 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 42-49, August 15, 1987.
- [2] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, Amsterdam: North-Holland, 1978.
- [3] J. H. Conway and N. J. A. Sloane, "Soft Decoding Techniques for Codes and Lattices," *IEEE Trans. Information Theory*, vol. IT-32, January 1986.
- [4] G. D. Forney, "Coset Codes I and II," *IEEE Trans. Information Theory*, to be published in 1988.

Table 1. Code words of (16,5) first-order Reed-Müller code and coset leaders

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	w_0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	w_1
0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1	w_2
1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0	w_3
0 0 0 0 1 1 1 1 0 0 0 0 1 1 1 1	w_4
1 1 1 1 0 0 0 0 1 1 1 1 0 0 0 0	w_5
0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1	w_6
1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0	w_7
0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1	w_8
1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0	w_9
0 0 0 0 1 1 1 1 1 1 1 1 0 0 0 0	w_{10}
1 1 1 1 0 0 0 0 0 0 0 0 1 1 1 1	w_{11}
0 0 1 1 0 0 1 1 1 1 0 0 1 1 0 0	w_{12}
1 1 0 0 1 1 0 0 0 0 1 1 0 0 1 1	w_{13}
0 0 1 1 1 1 0 0 0 0 1 1 1 1 0 0	w_{14}
1 1 0 0 0 0 1 1 1 1 0 0 0 0 1 1	w_{15}
1 1 0 0 0 0 1 1 0 0 1 1 1 1 0 0	w_{16}
0 0 1 1 1 1 0 0 1 1 0 0 0 0 1 1	w_{17}
0 1 0 1 0 1 0 1 1 0 1 0 1 0 1 0	w_{18}
1 0 1 0 1 0 1 0 0 1 0 1 0 1 0 1	w_{19}
0 1 0 1 1 0 1 0 0 1 0 1 1 0 1 0	w_{20}
1 0 1 0 0 1 0 1 1 0 1 0 0 1 0 1	w_{21}
1 0 1 0 0 1 0 1 0 1 0 1 1 0 1 0	w_{22}
0 1 0 1 1 0 1 0 1 0 1 0 0 1 0 1	w_{23}
0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0	w_{24}
1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1	w_{25}
0 1 1 0 0 1 1 0 1 0 0 1 1 0 0 1	w_{26}
1 0 0 1 1 0 0 1 0 1 1 0 0 1 1 0	w_{27}
1 0 0 1 0 1 1 0 1 0 0 1 0 1 1 0	w_{28}
0 1 1 0 1 0 0 1 0 1 1 0 1 0 0 1	w_{29}
1 0 0 1 0 1 1 0 0 1 1 0 1 0 0 1	w_{30}
0 1 1 0 1 0 0 1 1 0 0 1 0 1 1 0	w_{31}
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	A_0
1 1 0 0 1 0 1 0 1 1 1 1 0 1 1 0	A_1
1 0 1 1 0 1 0 0 1 0 1 1 1 0 1 1	A_2
1 1 1 0 0 1 1 1 0 1 0 0 1 1 0 1	A_3
0 1 1 0 1 1 0 0 0 0 0 0 0 1 0 1	B_0
0 1 0 1 0 1 1 0 1 1 1 1 0 0 1 1	B_1
1 0 0 0 1 1 0 1 1 0 1 1 1 1 1 0	B_2
1 1 0 1 0 0 0 1 0 1 0 0 1 0 0 0	B_3

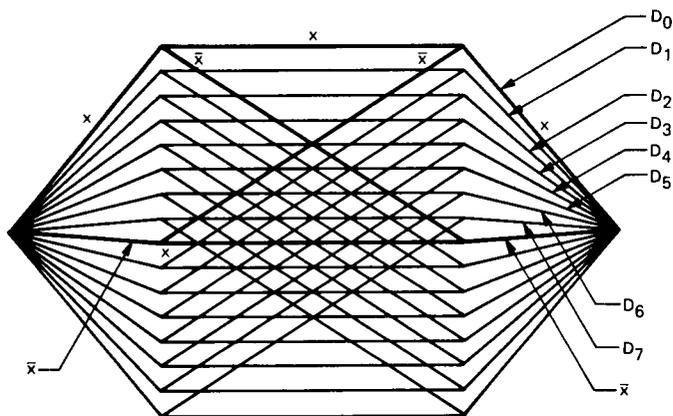


Fig. 1. The (24,5) subcode, $x \in A^*$

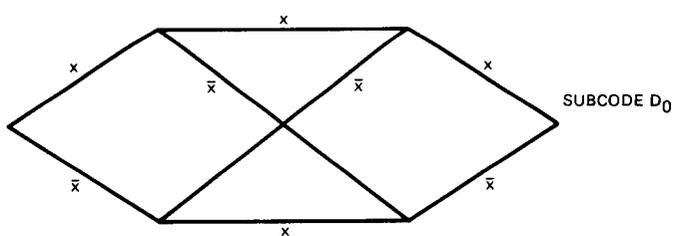


Fig. 2. The (24,2) subcode, $x = (0,0,0,0,0,0,0,0)$

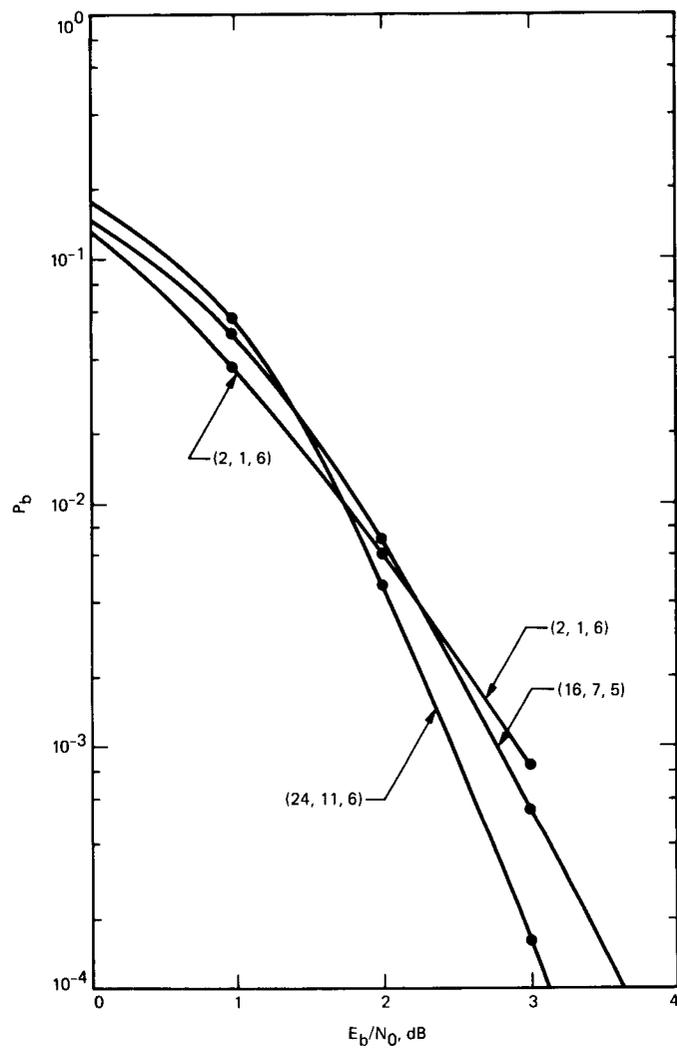


Fig. 3. Bit error probability of two new FS codes

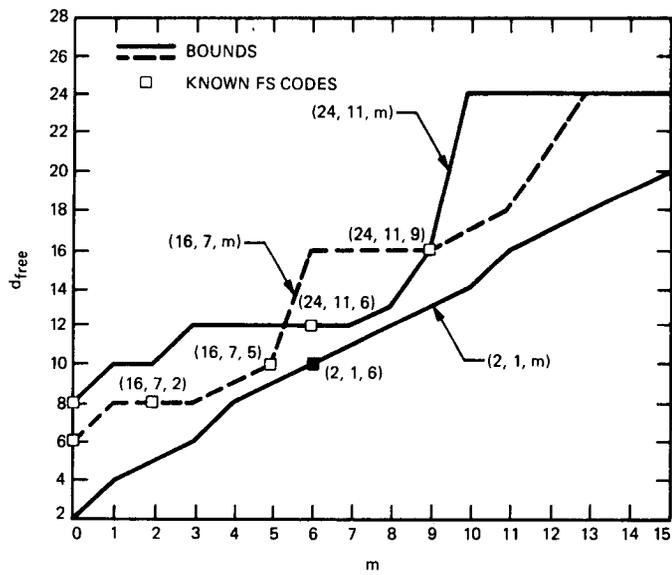


Fig. 4. Free distance bounds for two classes of FS codes compared to $(2,1,m)$ convolutional codes

A Comparison of VLSI Architectures for Time and Transform Domain Decoding of Reed-Solomon Codes

I. S. Hsu, T. K. Truong, L. J. Deutsch, and E. H. Satorius
Communications Systems Research Section

I. S. Reed
University of Southern California

It is well known that the Euclidean algorithm or its equivalent, continued fractions, can be used to find the error locator polynomial needed to decode a Reed-Solomon (RS) code. It is shown in this article that this algorithm can be used for both time and transform domain decoding by replacing its initial conditions with the Forney syndromes and the erasure locator polynomial. By this means both the error locator polynomial and the error evaluator polynomial can be obtained with the Euclidean algorithm.

With these ideas, both time and transform domain Reed-Solomon decoders for correcting errors and erasures are simplified and compared in this article. As a consequence, the architectures of Reed-Solomon decoders for correcting both errors and erasures can be made more modular, regular, simple, and naturally suitable for VLSI implementation.

I. Introduction

The Euclidean algorithm for solving the key equation for decoding Bose-Chaudhuri-Hocquenghem (BCH) and Goppa type codes was first developed by Sugiyama *et al.* [1]. The authors [2], [3] derived a fast decoding of Reed-Solomon (RS) codes using the continued fraction, which is closely related to the Euclidean algorithm. Brent and Kung [4] were the first to suggest a systolic array architecture for computing the greatest common divisor (GCD) of two polynomials. Through the use of these ideas, a pipeline structure for a transform domain decoder was developed to decode errors of RS codes [5]. An important ingredient of this design is a modi-

fied Euclidean algorithm for computing the error locator polynomial.

The computation of inverse field elements is completely avoided in the above-mentioned modification of Euclid's algorithm. Recently, the authors [6] proposed that a recursive algorithm could be used to perform this modified Euclidean algorithm. An important advantage of this new recursive algorithm is that the entire systolic array needed to perform Euclid's algorithm requires substantially less silicon area than the pipeline version of the modified Euclidean algorithm given in [5].

Forney [13] defined an errata locator polynomial using what are now called Forney syndromes to correct both errors and erasures. Blahut [7] showed that the errata locator polynomial can be computed directly by initializing Berlekamp's algorithm with the erasure locator polynomial.

Recently Eastman [8] suggested that the errata evaluator polynomial can be computed directly by initializing Berlekamp's algorithm with the Forney syndrome polynomial. This new, simplified decoding procedure is proved in [9]. By this technique, it is possible to compute the errata locator polynomial and the errata evaluator polynomial simultaneously from the Euclidean algorithm. This new RS decoder uses both the erasure locator polynomial and the Forney syndrome polynomial as initial conditions for the Euclidean algorithm.

It is shown and proved in [9] that the modified Euclidean algorithm mentioned above can be used to solve the Berlekamp-Massey key equation for the errata locator polynomial and the errata evaluator polynomial directly and simultaneously. By this means a new, simplified pipeline architecture for both the time and transform domain decoders can be developed for correcting both errors and erasures of RS codes. Such a decoding technique can be faster and simpler than previous methods [15], [10].

In this article, it is found that the VLSI implementation of the transform domain decoder is simpler than that of the time domain decoder. However, for a long RS code (10 bits or larger), due to the large size of the inverse transform unit needed in the transform decoder, the VLSI area needed to implement the transform domain decoder can be substantially larger than that needed for the time domain decoder. For moderately long codes, such as the 8-bit (255, 223) RS code used in the concatenated coding system for NASA's Voyager mission [11], the transform domain decoder is still simpler than the time domain decoder.

The above-mentioned NASA coding system is called the "baseline" system. It uses a (7, 1/2) convolutional code as its inner code and an 8-bit (255, 223) RS code as its outer code. It is shown [12] that this system achieves a bit-error rate (BER) of 10^{-6} at a bit signal-to-noise ratio (SNR) of 2.53 dB.

As mentioned above, the time domain decoder is more efficient in area than the transform domain decoder for very long RS codes. One such example is the long, 10 bits/symbol (1023, 959) RS code presently being considered for very deep space probes. If this code is concatenated with a (15, 1/5) convolutional code, it achieves a BER of 10^{-6} at an SNR of 0.5 dB [12]. Evidently the new NASA concatenated coding system provides a 2 dB improvement over the present baseline

system. It is for this reason and many other applications that it is important to develop an efficient, VLSI implementable, time domain RS decoder.

II. The Time Domain Decoder for RS Codes

An algorithm is developed in [15] for time domain decoding of RS codes to correct both errors and erasures through the use of continued fractions or their equivalent, Euclid's algorithm. This algorithm is a modification of the Berlekamp-Forney method [13], [14]. In this algorithm, the continued fraction algorithm is used to find the error locator polynomial from the remainder of the formal power series for the Forney syndrome. The disadvantage of this algorithm is that after the error locator polynomial is obtained by continued fractions, two polynomial multiplications are needed to compute the errata locator polynomial and the errata evaluator polynomial from the known error locator polynomial.

In this section, the above-mentioned algorithm is modified to correct both errors and erasures in the time domain decoding of RS codes by the use of the Euclidean algorithm. In this new algorithm, the Euclidean algorithm is used to solve the Berlekamp-Forney key equation for the errata locator polynomial and the errata evaluator polynomial directly and simultaneously. The advantage of this algorithm over previous methods [15] is that separate computation of the errata locator polynomial and the errata evaluator polynomial, which is usually needed [15], can be avoided. This new decoding algorithm is highly suitable for both VLSI and software implementation.

First, let $GF(2^m)$ be a finite field of 2^m elements. Also, let $N = 2^m - 1$ be the length of the (N, I) RS code over $GF(2^m)$ with minimum distance d , where $I = N - (d - 1)$ denotes the number of m -bit message symbols and $d - 1$ denotes the number of parity symbols such that $d - 1$ is either an even or an odd integer. The following five vectors are defined as:

$$\mathbf{c} = (c_0, c_1, \dots, c_{N-1}), \text{ code vector}$$

$$\mathbf{r} = (r_0, r_1, \dots, r_{N-1}), \text{ received vector}$$

$$\mathbf{e} = (e_0, e_1, \dots, e_{N-1}), \text{ error vector}$$

$$\mathbf{u} = (u_0, u_1, \dots, u_{N-1}), \text{ erasure vector}$$

$$\bar{\mathbf{u}} = (\bar{u}_0, \bar{u}_1, \dots, \bar{u}_{N-1}), \text{ errata vector}$$

These vectors are related by $\bar{\mathbf{u}} = \mathbf{e} + \mathbf{u}$ and $\mathbf{r} = \mathbf{c} + \mathbf{u} + \mathbf{e}$.

Suppose that t errors and v erasures occur in the received vector \mathbf{r} , and assume that $v + 2t \leq d - 1$. Next let α be a primitive element in $GF(2^m)$. Then $\gamma = \alpha^i$ is also a primitive element in $GF(2^m)$, where $(i, N) = 1$.

To minimize the complexity of an RS encoder it is desirable that the generator polynomial be symmetric. If γ is a root of the code's generator polynomial, it is shown [16] that the generator polynomial $g(x)$ is symmetric if and only if

$$g(x) = \prod_{i=b}^{b+(d-2)} (x - \gamma^i) = \sum_{i=0}^{d-1} g_i x^i \quad (1)$$

where $g_0 = g_{d-1} = 1$ and b satisfies the equality $2b + d - 2 = 2^m - 1$. The syndromes of the code are given by

$$\begin{aligned} S_{(b-1)+k} &= \sum_{i=0}^{N-1} \bar{u}_i \gamma^{i(b-1+k)} = \sum_{i=0}^{N-1} (u_i + e_i) \gamma^{i(b-1+k)} \\ &= \sum_{j=1}^{v+r} Y_j X_j^{(b-1)+k} \quad \text{for } 1 \leq k \leq d-1 \end{aligned} \quad (2)$$

where X_j is either the j th erasure or error location, and Y_j is either the j th erasure or the error magnitude. Define the set $\Lambda = \{X_i | X_i \text{ is an erasure location}\}$ and $\lambda = \{X_i | X_i \text{ is an error location}\}$. Define the syndrome polynomial

$$S(x) = \sum_{k=1}^{d-1} S_{(b-1)+k} x^{k-1} \quad (3a)$$

Then it is not difficult to show (see [14]) that

$$\begin{aligned} S(x) &= \sum_{k=1}^{d-1} S_{(b-1)+k} x^{k-1} = \sum_{j=1}^{v+r} \frac{Y_j X_j^b}{(1 - X_j x)} \\ &\quad - \sum_{j=1}^{v+r} \frac{Y_j X_j x^{b+d-1}}{(1 - X_j x)} \end{aligned} \quad (3b)$$

Following [14], we define four different polynomials as follows:

The erasure locator:

$$\Lambda(x) = \prod_{X_j \in \Lambda} (1 - X_j x) = \sum_{j=1}^v (1 - X_j x) = \sum_{j=0}^v (-1)^j \Lambda_j x^j \quad (4a)$$

where $\Lambda_0 = 1$.

The error locator:

$$\lambda(x) = \prod_{X_j \in \lambda} (1 - X_j x) = \sum_{j=1}^r (1 - X_j x) = \sum_{j=0}^r (-1)^j \lambda_j x^j \quad (4b)$$

where $\lambda_0 = 1$.

The errata locator:

$$\tau(x) = \Lambda(x)\lambda(x) = \prod_{j=1}^{v+r} (1 - X_j x) = \sum_{j=0}^{v+r} (-1)^j \tau_j x^j \quad (4c)$$

where $\tau_0 = 1$.

The errata evaluator:

$$A(x) = \sum_{j=1}^{v+r} Y_j X_j^b \left(\prod_{i \neq j} (1 - X_i x) \right) \quad (4d)$$

In terms of the polynomials defined above, Eq. (3b) becomes

$$S(x)\tau(x) = A(x) + x^{d-1} \sum_{j=1}^{v+r} Y_j X_j^{b+d-1} \left(\prod_{i \neq j} (1 - X_i x) \right) \quad (5)$$

From Eq. (5), one obtains the congruence relation,

$$S(x)\tau(x) \equiv A(x) \pmod{x^{d-1}} \quad (6a)$$

It is shown [9] that Eq. (6a) can be solved to yield

$$S(x) \equiv \frac{A(x)}{\lambda(x)\Lambda(x)} \pmod{x^{d-1}} \quad (6b)$$

It is well known, e.g., see [15], that the maximum number of errors in an RS code which can be corrected is $\lfloor (d-1-v)/2 \rfloor$ where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x , i.e., the principal part of x . We now define the Forney syndrome polynomial.

Definition 1: The Forney syndrome polynomial is defined by

$$T(x) \equiv S(x)\Lambda(x) \pmod{x^{d-1}} \quad (7)$$

By Eq. (7), the key in Eq. (6b) for $\lambda(x)$ and $A(x)$ is:

$$T(x) \equiv \frac{A(x)}{\lambda(x)} \bmod x^{d-1} \quad (8)$$

where

$$\deg \{\lambda(x)\} \leq \lfloor (d-1-\nu)/2 \rfloor$$

and

$$\deg \{A(x)\} \leq \lfloor (d+\nu-3)/2 \rfloor$$

It is shown in the following theorem that the errata evaluator polynomial $A(x)$ and the errata locator polynomial $\tau(x)$ can be obtained simultaneously and simply from knowing $T(x)$ in Eq. (7) and the new key equation in Eq. (8), which takes into account both errors and erasures.

Theorem 1: Let $T(x)$ in Eq. (7) be the Forney syndrome polynomial of a t -error and ν -erasure correcting RS code under the condition $\nu + 2t \leq d - 1$ where $d - 1$ is either an even or an odd integer. Consider the two polynomials $M(x) = x^{d-1}$ and $T(x) \equiv S(x) \Lambda(x) \bmod x^{d-1}$. Then the Euclidean algorithm for polynomials on $GF(2^m)$ can be used to develop two finite sequences $R_s(x)$ and $\tau_s(x)$ from the following two recursive formulas:

$$\tau_s(x) = (-q_{s-1}(x)) \tau_{s-1}(x) + \tau_{s-2}(x) \quad (9a)$$

and

$$R_s(x) = R_{s-2}(x) - q_{s-1}(x) R_{s-1}(x) \quad (9b)$$

for $(s = 1, 2, \dots)$, where the initial conditions are $\tau_0(x) = \Lambda(x)$, $\tau_{-1}(x) = 0$, $R_{-1}(x) = M(x)$, and $R_0(x) = T(x)$. Here $q_{s-1}(x)$ is obtained as the principal part of $R_{s-2}(x)/R_{s-1}(x)$. The recursion in Eq. (9) for $R_s(x)$ and $\tau_s(x)$ terminates when $\deg \{R_s(x)\} \leq \nu + w - 1$ for the first time for some value $s = s'$. Let

$$A(x) = \frac{R_{s'}(x)}{\Delta} \quad (10a)$$

and

$$\tau(x) = \frac{\tau_{s'}(x)}{\Delta} \quad (10b)$$

Also in Eq. (10), $\Delta = \tau_{s'}(0)$ is a field element in $GF(2^m)$ which is chosen so that $\tau_0 = 1$. Then $A(x)$ and $\tau(x)$ in Eq. (10) are the unique solutions of

$$A(x) \equiv T(x) \tau(x) \bmod x^{d-1} \quad (10c)$$

where both the inequalities, $\deg \{\tau(x)\} \leq \lfloor (d+\nu-1)/2 \rfloor$ and $\deg \{A(x)\} \leq \lfloor (d+\nu-3)/2 \rfloor$, are satisfied.

Proof: Theorem 1 is a proof [9] that the idea in [8] is correct.

The roots of $\tau(x)$ are the inverse locations of the t errors and ν erasures. These roots are most efficiently found by the Chien search procedure. By Eq. (4d) it is readily shown that the errata values are

$$Y_k = \frac{A(X_k)^{-1}}{(X_k^{d-1} \tau' X_k^{-1})} \quad \text{for } 1 \leq k \leq \nu + t \quad (11)$$

where $\tau'(X_k^{-1})$ is the derivative with respect to x of $\tau(x)$, evaluated at $x = X_k^{-1}$.

The overall time domain decoding of RS codes for correcting errors and erasures using the Euclidean algorithm is summarized in the following steps:

- (1) Compute the transform of the received m -tuple vector over $GF(2^m)$ from Eq. (2). Next, calculate the erasure locator polynomial $\Lambda(x)$ from Eq. (4a) and define $\deg \{\Lambda(x)\} = \nu$.
- (2) Compute the Forney syndrome polynomial from $T(x)$ in Eq. (7).
- (3) Determine the errata locator polynomial $\tau(x)$ and errata evaluator polynomial $A(x)$, where $0 \leq \nu < d - 1$, by applying the Euclidean algorithm to x^{d-1} and $T(x)$ as given by Eq. (7). The initial values of the Euclidean algorithm are $\tau_0(x) = \Lambda(x)$, $\tau_{-1}(x) = 0$, $R_{-1}(x) = x^{d-1}$, and $R_0(x) = T(x)$. The recursion in Eq. (9) for $R_s(x)$ and $\tau_s(x)$ terminates when $\deg \{R_s(x)\} \leq \lfloor (d+\nu-3)/2 \rfloor$ for the first time for some value $s = s'$. Finally, compute $\tau(x)$ and $A(x)$ from Eq. (10). For $\nu = d - 1$, set $\tau(x) = \Lambda(x)$ and $A(x) = T(x)$.
- (4) Compute the errata values from Eq. (11).

To illustrate the time domain decoding procedure for correcting errors and erasures, an elementary example of an RS code over $GF(2^4)$ is now presented. The representation of the field $GF(2^4)$ generated by the primitive irreducible polynomial $g(x) = x^4 + x + 1$ is given in Appendix A.

Example 1: Consider a (15, 9) RS code over $GF(2^4)$ with minimum distance $d = 7$. In this code, ν erasures and t errors under the condition $2t + \nu \leq d - 1$ can be corrected. In order

to simplify this example, let $\gamma = \alpha$ and $b = 1$. Thus, the generator polynomial of such a (15, 9) RS code is defined by

$$g(x) = \prod_{i=1}^6 (x - \alpha^i) = x^6 + \alpha^{10}x^5 + \alpha^{14}x^4 + \alpha^4x^3 + \alpha^6x^2 + \alpha^9x + \alpha^6$$

Assume the message symbols are

$$I(x) = \alpha^{10}x^{14} + \alpha^{12}x^{13} + \alpha^8x^{12} + \alpha^5x^{11} + \alpha^6x^{10} + \alpha^{14}x^9 + \alpha^{13}x^8 + \alpha^{11}x^7 + \alpha^9x^6$$

The encoded code word, which is a multiple of $g(x)$, is

$$c(x) = \alpha^{10}x^{14} + \alpha^{12}x^{13} + \alpha^8x^{12} + \alpha^5x^{11} + \alpha^6x^{10} + \alpha^{14}x^9 + \alpha^{13}x^8 + \alpha^{11}x^7 + \alpha^9x^6 + x^5 + \alpha x^4 + \alpha^2x^3 + \alpha^6x^2 + \alpha^{12}x + \alpha^8$$

Written as a vector, the code word is

$$\mathbf{c} = (\alpha^{10}, \alpha^{12}, \alpha^8, \alpha^5, \alpha^6, \alpha^{14}, \alpha^{13}, \alpha^{11}, \alpha^9, \alpha^0, \alpha, \alpha^2, \alpha^6, \alpha^{12}, \alpha^8)$$

Assume the erasure vector is

$$\mathbf{u} = (0, 0, 0, 0, 0, 0, 0, \alpha^2, 0, 0, 0, 0, 0, 0, 0) \quad (12)$$

and the error vector is

$$\mathbf{e} = (0, 0, 0, 0, \alpha^{11}, 0, 0, 0, 0, 0, 0, \alpha^7, 0, 0, 0) \quad (13)$$

Then the errata vector is

$$\bar{\mathbf{u}} = \mathbf{u} + \mathbf{e} = (0, 0, 0, 0, \alpha^{11}, 0, 0, \alpha^2, 0, 0, 0, \alpha^7, 0, 0, 0) \quad (14)$$

Assume the received vector is

$$\mathbf{r} = \mathbf{c} + \bar{\mathbf{u}} = (\alpha^{10}, \alpha^{12}, \alpha^8, \alpha^5, \alpha, \alpha^{14}, \alpha^{13}, \alpha^9, \alpha^9, \alpha^0, \alpha, \alpha^{12}, \alpha^6, \alpha^{12}, \alpha^8) \quad (15)$$

The syndromes S_k for r are

$$S_k = \sum_{n=0}^{14} r_n \alpha^{nk} = \alpha^7 (\alpha^3)^k + \alpha^2 (\alpha^7)^k + \alpha^{11} (\alpha^{10})^k$$

for $1 \leq k \leq 6$

This yields $S_1 = \alpha^0$, $S_2 = \alpha^{13}$, $S_3 = \alpha^{14}$, $S_4 = \alpha^{11}$, $S_5 = \alpha$, and $S_6 = 0$. Thus, the syndrome polynomial is $S(x) = \alpha^0 + \alpha^{13}x + \alpha^{14}x^2 + \alpha^{11}x^3 + \alpha x^4 + 0x^5$.

The erasure locator polynomial is $\Lambda(x) = (1 + \alpha^7x)$. In this example, the maximum erasure correcting capability is

$$\lfloor (d - 1 - \nu)/2 \rfloor = \lfloor (7 - 1 - 1)/2 \rfloor = 2$$

By Eq. (7), one obtains the Forney syndrome polynomial as

$$\begin{aligned} T(x) &\equiv \Lambda(x)S(x) \equiv (1 + \alpha^7x)(1 + \alpha^{13}x + \alpha^{14}x^2 + \alpha^{11}x^3 + \alpha x^4 + 0x^5) \pmod{x^6} \\ &\equiv (0x^6 + \alpha^8x^5 + \alpha^9x^4 + \alpha x^3 + \alpha^{12}x^2 + \alpha^5x + \alpha^0) \pmod{x^6} \\ &= \alpha^8x^5 + \alpha^9x^4 + \alpha x^3 + \alpha^{12}x^2 + \alpha^5x + \alpha^0 \end{aligned} \quad (16)$$

In Eq. (16), the coefficients of $T(x)$, $T_0 = \alpha^0$, $T_1 = \alpha^5$, $T_2 = \alpha^{12}$, $T_3 = \alpha$, $T_4 = \alpha^9$, and $T_5 = \alpha^8$ are the Forney syndromes.

The Euclidean algorithm is applied next to polynomial x^{d-1} and $T(x)$ in Eq. (16). By this means, polynomials $\tau(x)$ and $A(x)$ are determined next by use of the Euclidean algorithm. This is accomplished by the recursive Eqs. (9a) and (9b) illustrated in Table 1, where initially $R_{-1}(x) = x^{d-1} = x^6$ and $R_0(x) = \alpha^8x^5 + \alpha^9x^4 + \alpha x^3 + \alpha^{12}x^2 + \alpha^5x + 1$. From Table 1, one observes that $\deg \{R'_s(x)\} = \deg \{R_2(x)\} = 2 \leq \lfloor (d + \nu - 3)/2 \rfloor = 2$. Thus, the computation terminates at $s' = 2$, and

$$R_2(x) = \alpha^7x^2 + \alpha x + \alpha^2 \quad (17a)$$

and

$$\tau_2(x) = \alpha^7x^3 + \alpha^{13}x^2 + \alpha^4x + \alpha^2 \quad (17b)$$

By Eqs. (10a) and (10b), one has

$$\tau(x) = \frac{1}{\alpha^2} \tau_2(x) = \alpha^5 x^3 + \alpha^{11} x^2 + \alpha^2 x + 1 \quad (18)$$

and

$$A(x) = \frac{1}{\alpha^2} R_2(x) = \alpha^5 x^2 + \alpha^{14} x + 1 \quad (19)$$

By using Chien's search, the roots of $\tau(x)$ constitute the set $\{\alpha^{-7}, \alpha^{-3}, \alpha^{-10}\}$. The derivative with respect to x of $\tau(x)$ in Eq. (18) is $\tau'(x) = \alpha^5 x^2 + \alpha^2$. Thus, the errata values are

$$Y_1 = \frac{AX_1^{-1}}{\tau'(X_1^{-1})} = \frac{A(\alpha^{-7})}{\tau'(\alpha^{-7})} = \frac{\alpha^5 (\alpha^{-7})^2 + \alpha^{14} (\alpha^{-7}) + 1}{\alpha^5 (\alpha^{-7})^2 + \alpha^2} = \alpha^2$$

$$Y_2 = \frac{AX_2^{-1}}{\tau'(X_2^{-1})} = \frac{\alpha^5 (\alpha^{-3}) + \alpha^{14} (\alpha^{-3}) + 1}{\alpha^5 (\alpha^{-3})^2 + \alpha^2} = \alpha^7$$

and

$$Y_3 = \frac{AX_3^{-1}}{\tau'(X_3^{-1})} = \frac{\alpha^5 (\alpha^{-10})^2 + \alpha^{14} (\alpha^{-10}) + 1}{\alpha^5 (\alpha^{-10})^2 + \alpha^2} = \alpha^{11}$$

III. The Transform Decoder for RS Codes

The transform decoder of Gore and Mandelbaum [17], [18] was developed further in [10] to correct both errors and erasures. This decoding procedure was based on the algorithm originally invented by Forney [13] (also see [10]). By the above-mentioned Euclidean algorithm, the transform domain decoding procedure in [10] can be simplified further.

By the same procedure used in the time domain decoder, one can obtain the errata locator polynomial given in Eq. (4c). Hence,

$$\begin{aligned} \tau(X_i^{-1}) &= 1 + (-1)\tau_1(X_i^{-1}) + (-1)^2\tau_2(X_i^{-1})^2 \\ &+ \dots + (-1)^{\nu+t}\tau_{\nu+t}(X_i^{-1})^{\nu+t} = 0 \\ &\text{for } 1 \leq i \leq \nu + t \end{aligned} \quad (20)$$

Multiplying Eq. (20) by $Y_i X_i^{(b-1)+k}$ yields

$$\begin{aligned} Y_i X_i^{(b-1)+k} - \tau_1 Y_i X_i^{(b-1)+k-1} \\ + \dots + (-1)^{\nu+t} \tau_{\nu+t} Y_i X_i^{(b-1)+k-(\nu+t)} = 0 \end{aligned} \quad (21)$$

Summing Eq. (21) over i for $1 \leq i \leq \nu + t$ produces

$$\begin{aligned} \sum_{i=1}^{\nu+t} Y_i X_i^{(b-1)+k} - \tau_1 \sum_{i=1}^{\nu+t} Y_i X_i^{(b-1)+k-1} \\ + \dots + (-1)^{\nu+t} \tau_{\nu+t} \sum_{i=1}^{\nu+t} Y_i X_i^{(b-1)+k-(\nu+t)} = 0 \end{aligned} \quad (22)$$

From Eq. (22), one has

$$\begin{aligned} S_{(b-1)+k} - \tau_1 S_{(b-1)+k-1} \\ + \dots + (-1)^{\nu+t} \tau_{\nu+t} S_{(b-1)+k-(\nu+t)} = 0 \end{aligned} \quad (23)$$

Hence, in general,

$$\begin{aligned} E_{(b-1)+k} - \tau_1 E_{(b-1)+k-1} \\ + \dots + (-1)^{\nu+t} \tau_{\nu+t} E_{(b-1)+k-(\nu+t)} = 0 \end{aligned} \quad \text{for } k \geq d \quad (24)$$

are the recursive equations for E_j , the transforms of the errata pattern, where initially $E_b = S_b, E_{b+1} = S_{b+1}, \dots, E_{b+d-2}$ are known from the prior syndrome calculation.

From Eq. (24), one obtains the rest of the transform of $\bar{\mathbf{u}}$, i.e., the S_ℓ for $0 \leq \ell \leq N-1$. The amplitude $\bar{\mathbf{u}}$ vector is found by taking the inverse transform over $GF(2^m)$ of $S_\ell, 0 \leq \ell \leq N-1$. Finally, the original m -tuple code vector can be obtained by subtracting $\bar{\mathbf{u}}$ from the received vector \mathbf{r} .

Let us now recapitulate the above transform decoding algorithm of RS codes for correcting both errors and erasures, using transforms over $GF(2^m)$ and the Euclidean algorithm. This procedure is composed of the following five steps:

- (1) Use step 1 in the time domain decoder.
- (2) Use step 2 in the time domain decoder.
- (3) Use step 3 in the time domain decoder.
- (4) Compute the rest of the transform of the errata vector by the use of Eq. (24).
- (5) Invert the transform to recover the errata vector using the fact that $S_0 = S_N$. Then obtain the corrected code vector.

To illustrate the transform domain decoder for correcting errors with erasures, the data for the (15, 9) RS code over $GF(2^4)$ used in Example 1 is again used.

Example 2: Consider the (15, 9) RS code over $GF(2^4)$ with $d = 7$. For this code, the erasure, error, errata, and received vectors are given by Eqs. (12), (13), (14), and (15), respectively. By Eq. (18), the errata locator polynomial is

$$\tau(x) = \tau_0 + \tau_1 x + \tau_2 x^2 + \tau_3 x^3 = 1 + \alpha^2 x + \alpha^{11} x^2 + \alpha^5 x^3$$

where $\tau_0 = 1$, $\tau_1 = \alpha^2$, $\tau_2 = \alpha^{11}$, and $\tau_3 = \alpha^5$.

By Eq. (23), the rest of the transform of the errata vector is

$$S_k = \alpha^2 S_{k-1} + \alpha^{11} S_{k-2} + \alpha^5 S_{k-3} \quad \text{for } 7 \leq 15 \quad (25)$$

That is, $S_7 = \alpha^{13}$, $S_8 = \alpha^{13}$, $S_9 = \alpha^7$, $S_{10} = \alpha^3$, $S_{11} = \alpha^5$, $S_{12} = \alpha^{13}$, $S_{13} = \alpha^5$, $S_{14} = \alpha^5$, and $S_0 = 1$.

The inverse transform of S_k is

$$\bar{u}_k = \sum_{n=0}^{15-1} S_n \alpha^{-nk} \quad \text{for } 0 \leq k \leq 14$$

The result is $\bar{u} = (0, 0, 0, 0, \alpha^{11}, 0, 0, \alpha^2, 0, 0, 0, \alpha^7, 0, 0, 0)$. The corrected code is thus

$$\begin{aligned} c = r - \bar{u} &= (\alpha^{10}, \alpha^{12}, \alpha^8, \alpha^5, \alpha, \alpha^{14}, \alpha^{13}, \alpha^9, \alpha^9, \alpha^0, \alpha, \\ &\quad \alpha^{12}, \alpha^6, \alpha^{12}, \alpha^8) \\ &= (0, 0, 0, 0, \alpha^{11}, 0, 0, \alpha^2, 0, 0, 0, \alpha^7, 0, 0, 0) \\ &= (\alpha^{10}, \alpha^{12}, \alpha^8, \alpha^5, \alpha^6, \alpha^{14}, \alpha^{13}, \alpha^{11}, \alpha^9, \alpha^0, \alpha, \\ &\quad \alpha^2, \alpha^6, \alpha^{12}, \alpha^8) \end{aligned}$$

IV. A Comparison of VLSI Architecture of the Transform Decoder and the Time Domain Decoder

The block diagram of a (255, 223) RS time domain decoder is depicted in Fig. 1. Figure 2 shows the block diagram of a (255, 223) RS transform domain decoder. Each block diagram can be separated into two parts, indicated by broken lines, as shown in both Figs. 1 and 2. The first part, labeled as "I" in both block diagrams, has similar VLSI architecture. The major

functional units in this part are (1) the syndrome computation unit; (2) the power calculation unit; (3) the power expansion unit; (4) the polynomial expansion unit; and (5) the $\lfloor (d + \nu - 3)/2 \rfloor$ generator. Also included in this part are some delay registers. The lengths of the delay registers may not be equal in these two decoder architectures, but since they contain only replicated register cells, they can be considered identical in architecture.

Figure 3 shows the block diagram of the syndrome computation unit. This unit accepts the received messages and computes their syndromes. There are 32 syndrome subcells in a (255, 223) RS decoder. Each subcell depicted in Fig. 3 performs the operation as $S_i \leftarrow (S_i + r_i \alpha^i)$, where " \leftarrow " denotes the operation "is replaced by." The Berlekamp multiplier is used in this syndrome unit due to its simplicity in VLSI design [19]. The computed syndrome polynomial is labeled as $S(x)$ in both Figs. 1 and 2. In the time domain and transform domain decoders, the coefficients of $S(x)$ are fed in parallel to the polynomial expansion unit to compute the Forney syndromes.

The power calculation unit converts the received 1's and 0's into a sequence of α^k 's and 0's, where α is a primitive element of the finite field over which the RS code is defined. These received 1's and 0's indicate the occurrence or nonoccurrence, respectively, of an erasure at a specific location. Figure 4 shows the block diagram of the power calculation unit. Since the maximum erasure correcting capability of a (255, 223) RS decoder is 32, only 32 symbol latches are needed to store the locations of all the correctable erasures.

A detection circuit for detecting the occurrence of erasures is included in the power calculation unit. If an erasure occurs at the k th location, its corresponding symbol α^k is calculated and latched. This α^k 's sequence is fed to the polynomial expansion circuit, to the power expansion unit, and to the $\lfloor (d + \nu - 3)/2 \rfloor$ generator.

The power expansion unit converts the α^k 's sequence into an erasure locator polynomial $\Lambda(x)$ which has α^k 's as its roots. Figure 5 depicts the block diagram of this unit. The erasure locator polynomial $\Lambda(x)$ is fed to the modified GCD unit as one of the initial conditions.

A generator is used to compute $\lfloor (d + \nu - 3)/2 \rfloor$. This is shown in both Figs. 1 and 2. The output of this generator is sent to the modified GCD unit and used as a stop indicator for Euclid's algorithm.

Figure 6 presents a block diagram of the polynomial expansion circuit. The Forney syndromes for either the time domain decoder or the transform decoder are calculated in this unit.

Figure 7 depicts the block diagram of the modified GCD unit. As described previously in [6], a multiplexing scheme can be applied to the modified GCD unit to reduce the number of cells needed. The polynomial $\Lambda(x)$ together with the Forney syndrome polynomial $T(x)$ are the two inputs to the modified GCD unit. The output of the modified GCD unit is the errata locator polynomial $\tau(x)$ and the errata evaluator polynomial $A(x)$. The error correcting capability of the code is computed by $\lfloor (32 - \nu)/2 \rfloor$.

The differing functional units of the time and transform domain decoders are shown in the second half of Figs. 1 and 2 and are labeled as "II." One output of the modified GCD unit of the time domain decoder, the errata locator polynomial $\tau(x)$, is fed to a Chien search unit and to another unit for computing $[x^{b-1}\tau'(x)]^{-1} = [x^{111}\tau'(x)]^{-1}$, where $b = 112$ in this design. The other output of the modified GCD unit of the time domain decoder, the errata evaluator polynomial $A(x)$, is fed to the polynomial evaluation unit to perform the evaluation of $A(x)$. Figure 8 shows the block diagram of the polynomial evaluation unit.

The $[x^{111}\tau'(x)]^{-1}$ unit performs the calculation of one part of the errata magnitude [6]. Figure 9 depicts the block diagram of this unit. The product of the outputs of the polynomial evaluation unit and the $[x^{111}\tau'(x)]^{-1}$ unit forms the errata magnitude.

In the time domain decoder, the Chien search unit is used to search for the error and erasure locations; for more details, see [6]. The architecture of the Chien search unit is similar to that of the polynomial evaluation unit, except there is a zero detector at the end.

On the other hand, for the transform domain decoder design, the output from the modified GCD unit is the errata locator polynomial $\tau(x)$. This output is fed to the transform error pattern unit, along with the syndromes from the syndrome computation unit, to calculate the extended syndromes. A new architecture for the transform of the error-pattern unit is developed in Appendix A. The realization of this idea is

shown in the block diagram of the transform of the error-pattern unit, given in Fig. 10.

The computation of extended syndromes, together with the original syndromes, is sent to the inverse transform unit to obtain the estimated error patterns. Figure 11 shows the block diagram of the inverse transform error-pattern unit. It is easy to see that the architecture for the inverse transform unit is similar to that of the syndrome computation unit except that 255 subcells are needed in the inverse transform unit while the syndrome computation unit needs 32 subcells.

Clearly, the architecture of the transform domain decoder design is simpler than that of the time domain decoder design. This is because the transform domain decoder design needs only two regular function blocks in part II of Fig. 2. However, the time domain decoder requires three function blocks for the implementation in part II of Fig. 1.

Furthermore, the inverse-transform unit in the transform domain design contains 255 similar cells in the (255, 223) RS decoder. It is estimated that these 255 cells occupy only a moderate amount of silicon area, and that their geometric arrangement can be regular and simple. Therefore, substantial time for the design and test of such a VLSI chip can be saved. However, the advantage of the transform domain decoder is valid only for moderately short length RS codes. If long length RS codes are used to enhance the system's performance [12], the transform domain decoder needs a large inverse transform block. This might cause a problem in the VLSI implementation. In general, if a $GF(2^m)$ field is used to define an RS code, an inverse transform block composed of $2^m - 1$ cells is needed. Hence, the number of cells needed in an inverse transform block increases exponentially with the integer m . However, the number of transistors needed in the time domain decoder goes up only linearly as the integer m increases. Therefore, for long length codes, the time domain decoder is the more appealing approach. Although the computation of the time domain decoder is more complex than that of the transform domain decoder, for long RS codes the number of transistors needed in a time domain decoder is substantially less than that in a transform domain decoder.

Acknowledgment

We would like to acknowledge the important suggestion of W. L. Eastman of the MITRE Corp. that both the errata evaluator and locator polynomials can be obtained directly by only a modification of the initial conditions of the Berlekamp-Massey algorithm or Euclid's algorithm.

References

- [1] Y. Sugiyama, M. Kasahara, S. Hirasawa, and T. Namekawa, "A Method for Solving Key Equation for Decoding Goppa Codes," *IEEE Trans. on Contr.*, vol. 27, pp. 87-99, 1975.
- [2] I. S. Reed, R. A. Scholtz, T. K. Truong, and L. R. Welch, "The Fast Decoding of Reed-Solomon Codes Using Fermat Theoretic Transforms and Continued Fractions," *IEEE Trans. on Information Theory*, vol. IT-24, no. 1, pp. 100-106, January 1978.
- [3] L. R. Welch and R. A. Scholtz, "Continued Fractions and Berlekamp's Algorithm," *IEEE Trans. on Information Theory*, vol. IT-25, pp. 19-27, 1979.
- [4] R. P. Brent and H. T. Kung, "Systolic VLSI Arrays for Polynomial GCD Computation," *IEEE Trans. on Computers*, vol. C-33, no. 8, pp. 731-736, August 1984.
- [5] H. M. Shao, T. K. Truong, L. J. Deutsch, J. H. Yuen, and I. S. Reed, "A VLSI Design of a Pipeline Reed-Solomon Decoder," *IEEE Trans. on Computers*, vol. C-34, no. 5, pp. 393-403, May 1985.
- [6] H. M. Shao, T. K. Truong, I. S. Hsu, L. J. Deutsch, and I. S. Reed, "A Single Chip VLSI Reed-Solomon Decoder," *TDA Progress Report 42-84*, vol. October-December 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 73-81, February 15, 1986.
- [7] R. E. Blahut, *Theory and Practice of Error Control Codes*, New York: Addison-Wesley, p. 258, May 1984.
- [8] W. L. Eastman, *Decoding Erasures*, Bedford, Massachusetts: Mitre Corporation, 1986.
- [9] T. K. Truong, W. L. Eastman, I. S. Reed, and I. S. Hsu, "A Simplified Procedure for Correcting Both Errors and Erasures of a Reed-Solomon Code Using the Euclidean Algorithm," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 200-212, November 15, 1987.
- [10] I. S. Reed, T. K. Truong, and R. L. Miller, "Simplified Algorithm for Correcting Both Errors and Erasures of Reed-Solomon Codes," *Proc. IEE*, vol. 126, no. 10, pp. 961-963, October 1979.
- [11] R. L. Miller, L. J. Deutsch, and S. A. Butman, *On the Error Statistics of Viterbi Decoding and the Performance of Concatenated Codes*, JPL Publication 81-9, Jet Propulsion Laboratory, Pasadena, California, September 1981.
- [12] J. H. Yuen and Q. D. Vo, "In Search of a 2-dB Coding Gain," *TDA Progress Report 42-83*, vol. July-September 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 26-33, November 15, 1985.
- [13] G. D. Forney, "On Decoding BCH Codes," *IEEE Trans. on Information Theory*, vol. IT-11, pp. 549-557, 1965.
- [14] E. R. Berlekamp, *Algebraic Coding Theory*, New York: McGraw-Hill, 1968.
- [15] I. S. Reed, T. K. Truong, and R. L. Miller, "Decoding of BCH and RS Codes With Errors and Erasures Using Continued Fractions," *Electronics Letters*, vol. 15, no. 17, pp. 542-544, August 16, 1976.
- [16] E. R. Berlekamp, "Bit-Serial Reed-Solomon Encoders," *IEEE Trans. on Information Theory*, vol. IT-28, no. 6, pp. 869-874, November 1982.

- [17] W. C. Gore, *Transmitting Binary Symbols With Reed-Solomon Code*, Johns Hopkins Electrical Engineering Department Report 73-75, April 1973.
- [18] D. Mandelbaum, "On Decoding Reed-Solomon Codes," *IEEE Trans. on Information Theory*, vol. IT-17, pp. 707-712, 1971.
- [19] I. S. Hsu, T. K. Truong, H. M. Shao, L. J. Deutsch, and I. S. Reed, "A Comparison of VLSI Architecture of Finite Field Multipliers Using Dual, Normal or Standard Bases," *TDA Progress Report 42-90*, vol. April-June 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 63-75, August 15, 1987.
- [20] L. Johnson, U. Weiser, D. Cohen, and A. Davis, "Towards a Formal Treatment of VLSI Arrays," presented at the Caltech Conference on VLSI, Pasadena, California, January 1981.

Table 1. An example of the Euclidean algorithm used to find $\tau(x)$ and $A(x)$

S	$R_{s-2}(x) = q_{s-1}(x)R_{s-1}(x) + R_s(x)$	$q_{s-1}(x)$	$R_s(x)$	$\tau_s(x)$
-1			x^6	0
0			$\alpha^8 x^5 + \alpha^9 x^4 + \alpha x^3 + \alpha^{12} x^2 + \alpha^5 x + 1$	$1 + \alpha^7 x$
1	$\frac{\alpha^8 x^5 + \alpha^9 x^4 + \alpha x^3 + \alpha^{12} x^2 + \alpha^5 x + 1}{\frac{1}{\alpha^8} x + \frac{\alpha}{\alpha^8}}$ $\frac{\alpha^8 x^5 + \alpha^9 x^4 + \alpha x^3 + \alpha^{12} x^2 + \alpha^5 x + 1}{x^6 + \alpha x^5 + \alpha^8 x^4 + \alpha^4 x^3 + \alpha^{12} x^2 + \alpha^7 x}$ $\frac{\alpha x^5 + \alpha^8 x^4 + \alpha^9 x^3 + \alpha^{12} x^2 + \alpha^7 x}{\alpha x^5 + \alpha^2 x^4 + \alpha^9 x^3 + \alpha^5 x^2 + \alpha^{13} x + \alpha^8}$ $\frac{\alpha x^4 + \alpha^{14} x^3 + \alpha^{14} x^3 + \alpha^3 x + \alpha^8}{\alpha^8 x + 1}$	$\frac{1}{\alpha^8}(x + \alpha)$	$x^4 + \alpha^{14} x^3 + \alpha^6 x^2 + \alpha^2 x + \alpha^8$	$\frac{1}{\alpha^8}(x + \alpha)(1 + \alpha^7 x)$ $+ 0 = \frac{1}{\alpha^8}(\alpha^7 x^2 + \alpha^2 x + 2)$
2	$\frac{\alpha^8 x + 1}{\alpha^8 x^5 + \alpha^9 x^4 + \alpha^5 x + \alpha^8}$ $\frac{\alpha^8 x^5 + \alpha^9 x^4 + \alpha^7 x^3 + \alpha^3 x^2 + \alpha x}{x^4 + \alpha^{14} x^3 + \alpha x^2 + \alpha^2 x + 1}$ $\frac{\alpha^8 x^5 + \alpha^9 x^4 + \alpha^7 x^3 + \alpha^3 x^2 + \alpha x}{x^4 + \alpha^{14} x^3 + \alpha^{14} x^2 + \alpha^5 x + \alpha^8}$ $\frac{\alpha^7 x^2 + \alpha x + \alpha^2}{\alpha^7 x^2 + \alpha x + \alpha^2}$	$\alpha^8 x + 1$	$\alpha^7 x^2 + \alpha x + \alpha^2$	$(\alpha^7 x^2 + \alpha^2 x + \alpha) \cdot (x + \alpha^7)$ $+ (1 + \alpha^7 x)$ $= \alpha^7 x^3 + \alpha^{13} x^2 + \alpha^4 x + \alpha^2$

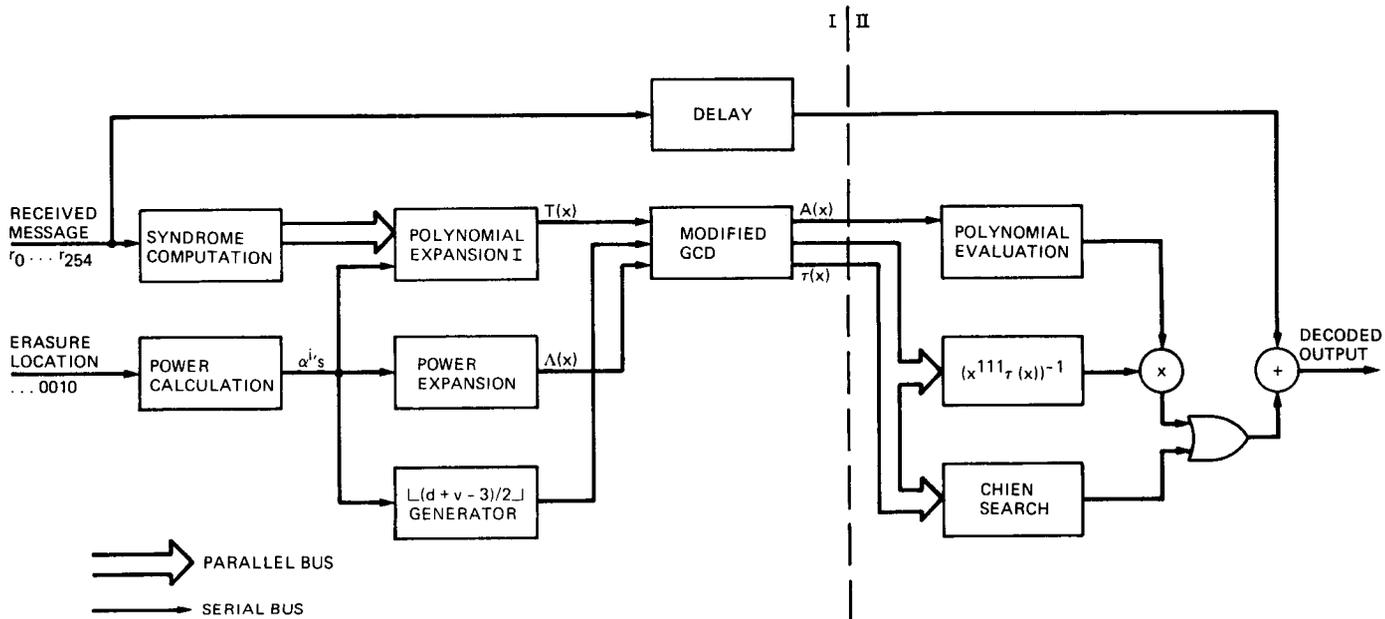


Fig. 1. Overall block diagram of a pipeline (255,223) RS time domain decoder

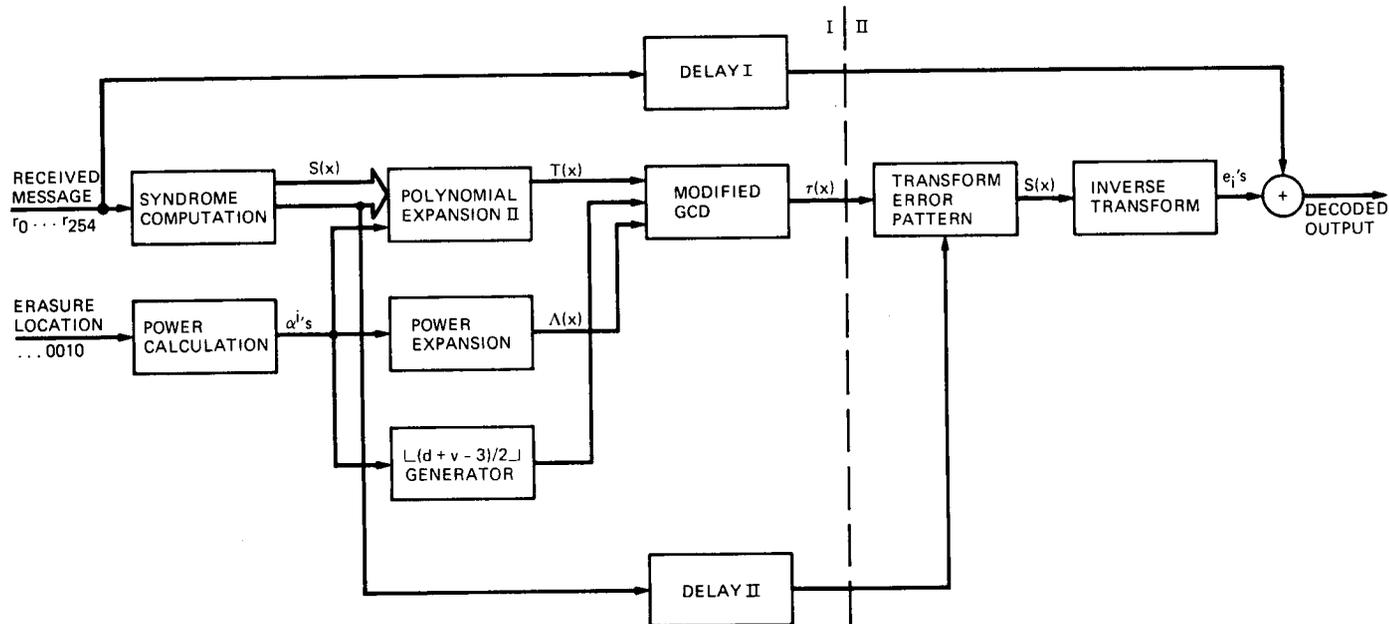


Fig. 2. Overall block diagram of a pipeline (255,223) RS transform domain decoder

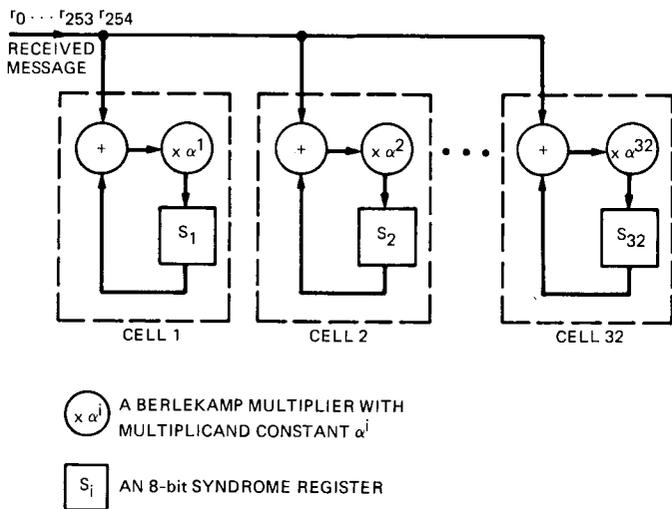


Fig. 3. Block diagram of the syndrome computing chip

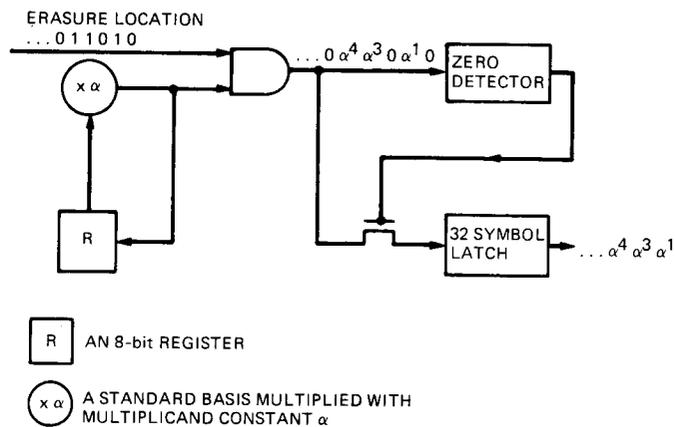


Fig. 4. Block diagram of the power calculation unit

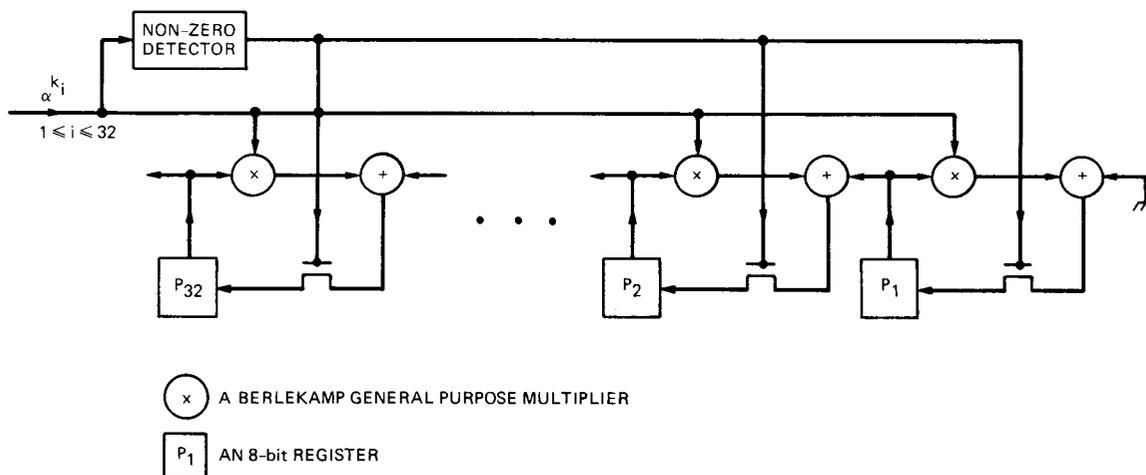


Fig. 5. Block diagram of the power expansion unit

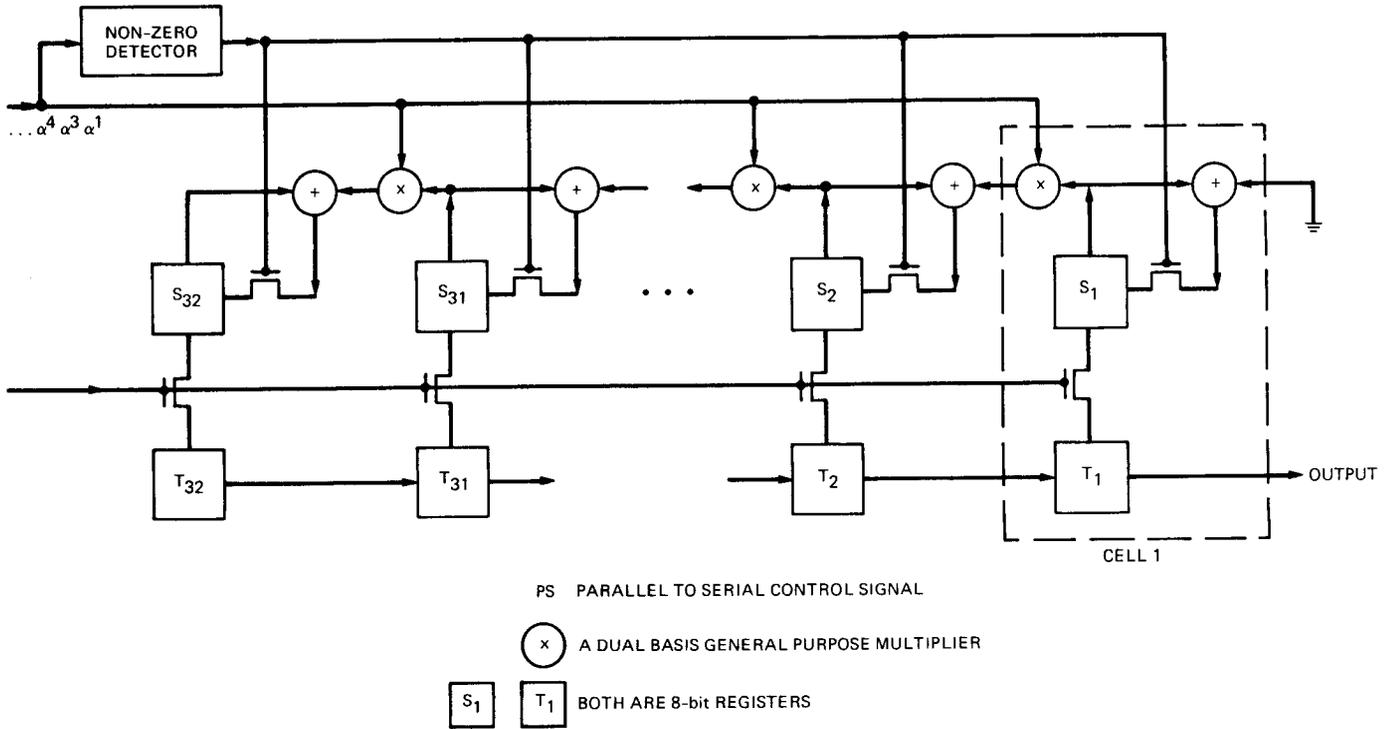


Fig. 6. Block diagram of a polynomial expansion unit

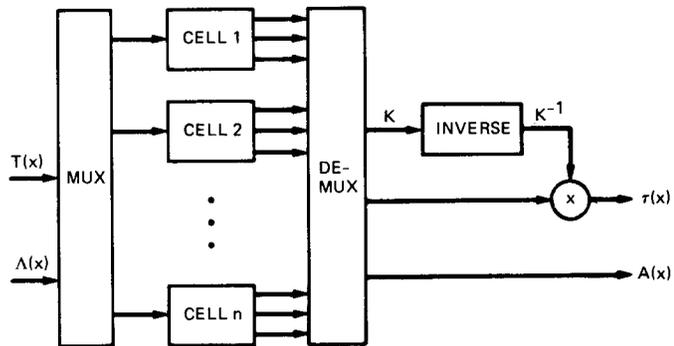
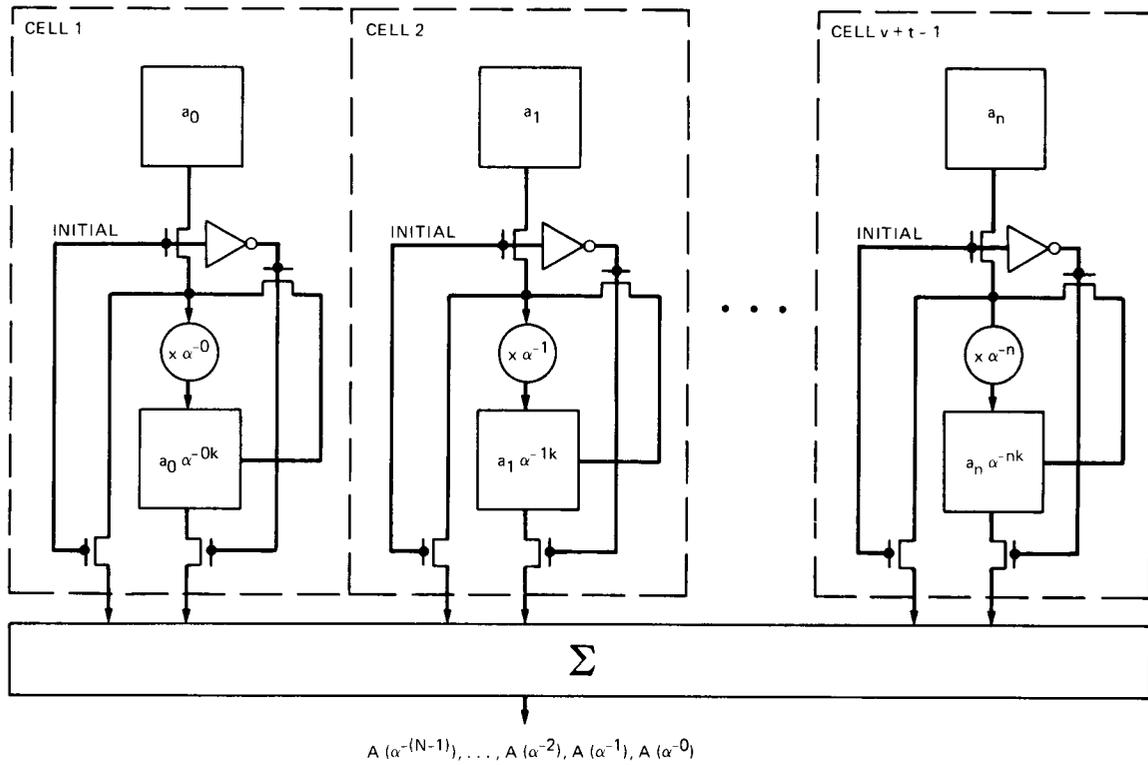


Fig. 7. Block diagram of the modified GCD unit



 A BERLEKAMP MULTIPLIER WITH MULTIPLICAND FIXED AS α^i  AN 8-bit REGISTER
 $n = v + t - 1$

Fig. 8. Block diagram of the polynomial evaluation unit

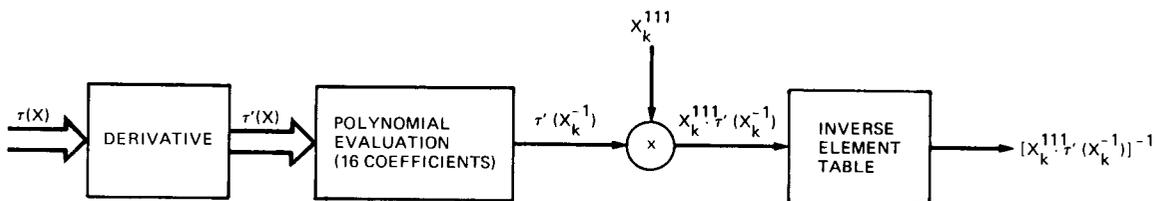


Fig. 9. Block diagram of the $[x^{111} r'(x)]^{-1}$ unit

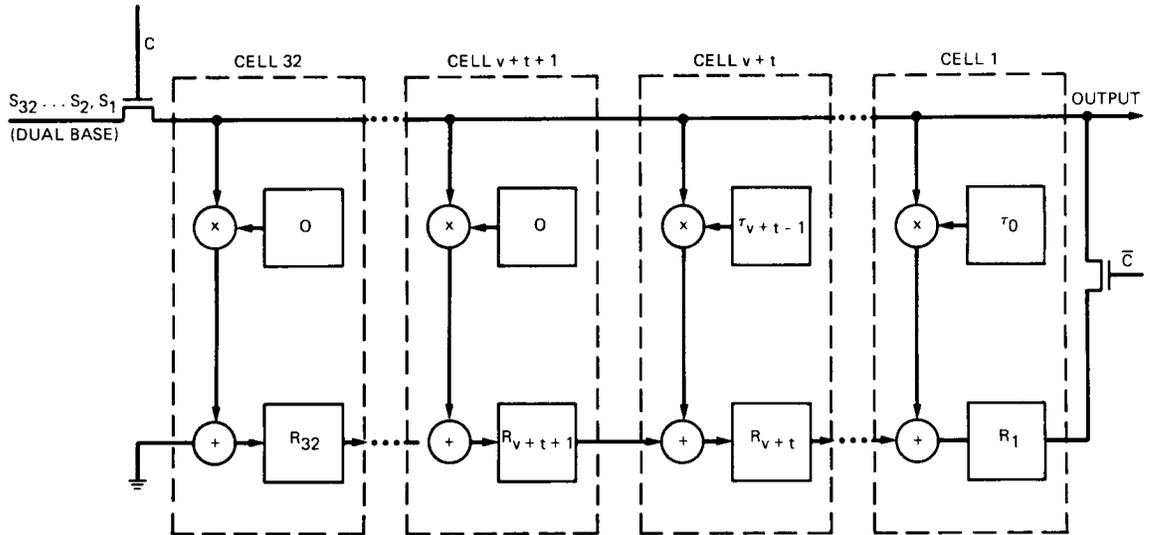


Fig. 10. Block diagram of the transform of the error pattern

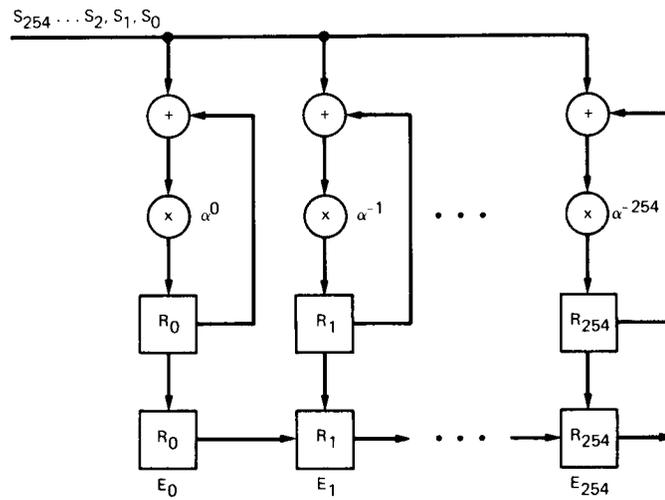


Fig. 11. Block diagram of the inverse transform unit

Appendix A

New Architecture for the Transform of the Error Pattern Unit

In this appendix, a VLSI architecture is developed to compute the transform of the error pattern. Recently, Johnson *et al.* [20] proposed a systolic array for computing a linear convolution. Using a technique similar to that suggested by Johnson *et al.*, the transform of the error pattern in Eq. (23) or Eq. (24) can be implemented in a systolic array. The advantage of this method over the previous method in Fig. 15 of [5] is that the long delay needed in the large XOR tree used for summing all the terms in Eq. (23) is eliminated. Also, the zero detectors needed in the previous design [5] are not required in this new architecture.

To illustrate this new architecture, the data in Example 2 for a (15, 9) RS code are used here as an example. The recursive equation to compute the remainder of the transform of the error pattern is given in Eq. (25). The new design for computing Eq. (25) is shown in Fig. A-1. In this figure, the function of each cell can be described by a register transfer

relation of the type $R_i \leftarrow R_{i+1} + S_k \alpha^i$. The input data are sent to all the cells simultaneously.

To understand the operation of this circuit, assume initially that all registers R_i for $1 \leq i \leq 3$ are set to zero. The control signal C is high for 6 symbol clocks to allow data S_1, S_2, \dots, S_6 to be fed into the circuit. The input data are also sent to the output node. At the same time, the complement signal \bar{C} of signal C is low to prevent the data stored in register R_1 from being sent to the output node. Note that one "clock time" for one Galois field symbol equals 4 circuit clock times. At the seventh symbol clock time the control signal C is switched to low or zero so that $\bar{C} = 1$. Therefore, the data stored in register R_1 , which equals S_7 at that moment, is sent to the output node and fed back to all basic cells. This process continues until the rest of the transform of the error pattern, i.e., S_7, S_8, \dots, S_{15} , is obtained. The detailed operation of this circuit is described in Table A-2.

Table A-1. Representations of the elements of $GF(2^4)$ generated by $\alpha^4 + \alpha + 1 = 0$

	α^3	α^2	α	α^0
α^0	0	0	0	1
α^1	0	0	1	0
α^2	0	1	0	0
α^3	1	0	0	0
α^4	0	0	1	1
α^5	0	1	1	0
α^6	1	1	0	0
α^7	1	0	1	1
α^8	0	1	0	1
α^9	1	0	1	0
α^{10}	0	1	1	1
α^{11}	1	1	1	0
α^{12}	1	1	1	1
α^{13}	1	1	0	1
α^{14}	1	0	0	1

Table A-2. The fifteen steps of the transform of the error pattern algorithm

Symbol clock time	R_1	R_2	R_3
1	$\alpha^5 S_1$	$\alpha^{11} S_1$	$\alpha^2 S_1$
2	$\alpha^5 S_2$	$\alpha^5 S_1 + \alpha^{11} S_2$	$\alpha^{11} S_1 + \alpha^2 S_2$
3	$\alpha^5 S_3$	$\alpha^5 S_2 + \alpha^{11} S_3$	$\alpha^2 S_3 + \alpha^{11} S_2 + \alpha^5 S_1 = S_4$
4	$\alpha^5 S_4$	$\alpha^{11} S_4 + \alpha^5 S_3$	$\alpha^2 S_4 + \alpha^{11} S_3 + \alpha^5 S_2 = S_5$
⋮	⋮	⋮	⋮
15	$\alpha^5 S_{14}$	$\alpha^{11} S_{14} + \alpha^5 S_{13}$	$\alpha^2 S_{14} + \alpha^{11} S_{13} + \alpha^5 S_{12} = S_{15} = S_0$

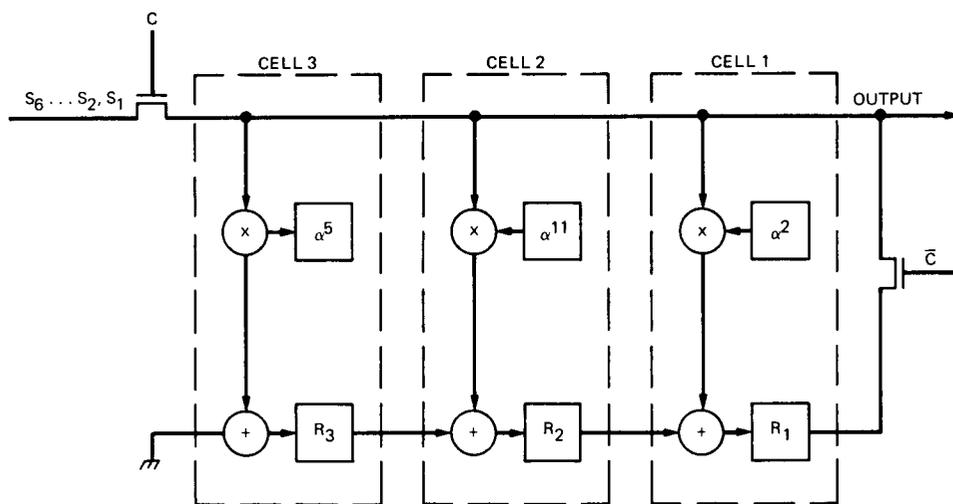


Fig. A-1. A systolic array for computing the transform of the error pattern in Eq. (25)

Spacecraft Telecommunications System Mass Estimates

J. H. Yuen

Telecommunications Science and Engineering Division

L. L. Sakamoto

Telecommunications Systems Section

Mass is the most important limiting parameter for present-day planetary spacecraft design. In fact, the entire spacecraft design can be characterized by mass. The more efficient the design of a spacecraft, the less mass will be required. The communications system is an essential and integral part of planetary spacecraft. In this article, a study is presented of the mass attributable to the communications system for spacecraft designs used in recent missions in an attempt to help guide future design considerations and research-and-development efforts.

The basic approach is to examine the spacecraft by subsystem and allocate a portion of each subsystem to telecommunications. Conceptually, this is to divide the spacecraft into two parts, telecommunications and non-telecommunications. In this way, it is clear what the mass attributable to the communications system is.

The percentage of mass is calculated using the actual masses of the spacecraft parts, except in the case of CRAF. In that case, estimated masses are used since the spacecraft has not been built. The results show that the portion of the spacecraft attributable to telecommunications is substantial. The mass fraction for Voyager, Galileo, and CRAF (Mariner Mark II) is 34 percent, 19 percent, and 18 percent, respectively. The large reduction of telecommunications mass from Voyager to Galileo is mainly due to the use of a deployable antenna instead of the solid antenna on Voyager.

I. Statement of the Problem

The task at hand is to separate the spacecraft into two parts: telecommunications and non-telecommunications. The Voyager spacecraft, for example, is made up of the 25 subsystems listed in Table 3.¹ Normally, the Radio Frequency Subsystem (RFS), the Modulation/Demodulation Subsystem (MDS), and the S/X-Band Antenna Subsystem (SXA) are said to compose the

telecommunications system, i.e., the communications system between the spacecraft and ground station. However, portions of many of the other subsystems are directly related or necessary to the support and function of these three subsystems. Therefore, a method must be created that will take those related portions of the other subsystems into account.

II. Analysis

The Voyager spacecraft is used here to illustrate the analysis. This method will be used to obtain results for the other spacecraft. We have already done so for the three chronologi-

¹Mariner Jupiter/Saturn 1977, Project Document 618-205, Functional Requirements Book, vols. 1 and 2 (internal document), Jet Propulsion Laboratory, Pasadena, California, 1977.

cally representative spacecraft Voyager, Galileo, and CRAF. To obtain overall percentages, each subsystem is analyzed individually. However, the same method of analysis cannot be used on all the subsystems. The detailed Voyager weight list is used for this calculation.²

A subjective analysis of the Voyager spacecraft telecommunications system mass was previously conducted by R. M. Dickinson.³ This resulted in a figure of fully one-third of the spacecraft mass. The present qualitative analysis, resulting in a figure of 34 percent, concurs with and verifies this conclusion.

The Radio Frequency Subsystem (RFS), Modulation/Demodulation Subsystem (MDS), and S/X-Band Antenna Subsystem (SXA) are the three telecommunications subsystems; therefore all (100 percent) of their mass is allocated to telecommunications.

The ten Science Instruments Subsystems (SCIs) are used for the purpose of gathering scientific data. This is clearly non-telecommunications, so none of the mass is allocated to telecommunications.

The Pyrotechnic Subsystem (PYRO) and the Systems Assembly Hardware (SAH) are clearly non-telecommunications, so none of the mass is allocated to telecommunications. The PYRO subsystem effects the launch vehicle/spacecraft separation, deployed booms, etc., while the SAH subsystem consists of parts needed to assemble the spacecraft.

The Flight Data Subsystem (FDS), Computer Command Subsystem (CCS), and Attitude and Articulation Control Subsystem (AACS) are the three on-board computers. Since these cannot really be broken into parts, the mass estimate is based on how much computing power was devoted to telecommunications. An estimate given by G. W. Garrison⁴ was approximately 10 percent of FDS, approximately 10 percent of CCS, and less than 5 percent of AACS. The total FDS mass (19.32 kg) includes the Reed-Solomon coder hardware (2.35 kg). This is considered to be telecommunications mass. Thus, 100 percent of 2.35 kg plus 10 percent of 16.97 kg (the total mass less the RS coder) yields 4.05 kg, or 20.95 percent of the FDS

mass allocated to telecommunications. Ten percent of the CCS mass and 5 percent of the AACS mass is allocated to telecommunications.

The Data Storage Subsystem (DSS) is the tape recorder and basically serves as a time buffer and provides functional redundancy for the telecommunications system. Since it wholly supports the telecommunications system, all (100 percent) is allocated to telecommunications.

The Structure Subsystem (STRU), Cabling Subsystem (CABL), Temperature Control Subsystem (TEMP), and Mechanical Devices Subsystem (DEV) are handled differently. For these four subsystems, the detailed equipment mass list is examined (see footnote 2), and each item is sorted into the categories of fully telecommunication, fully non-telecommunication, and partly telecommunication. The "partly telecommunication" category is further subdivided into six categories: PWR (Power), CCS, FDS, AACS, PROP (Propulsion), and DSS, depending on which subsystem they are related to. Fully telecommunication includes any items related to the RFS, MDS, and SXA subsystems, such as the High Gain Antenna (HGA). Fully non-telecommunication includes those items related to the SCI subsystems such as the scan platform and magnetometer boom. This category also contains such miscellaneous items as the phonograph record.

For each of the four subsystems, a percentage of the mass in each of the eight categories is taken and then summed together. One hundred percent is taken for fully telecommunication, 0 percent for fully non-telecommunication, and for the partly telecommunication categories, 10 percent of CCS and FDS, 5 percent of AACS, and 100 percent of DSS. The percentage for PWR (40 percent for Voyager) is taken to be the percentage of telecommunication power as derived in Appendix A. The percentage for PROP is taken to be the percentage of telecommunication mass (34 percent for Voyager) since almost all of the propellant is allocated to trajectory correction maneuvers. Only a very small percentage of the propellant (0.5 percent in the case of Galileo) is allocated to keeping the High Gain Antenna pointed toward the Earth (because of a different engine design, this may be more significant for CRAF). To illustrate, see Table 1 for the calculation of STRU. The complete mass calculation is shown in Table 4. The results of the Voyager mass calculation can be seen in Table 2.

III. Results and Conclusions

We have examined the mass of the telecommunications systems of three representative spacecraft: Voyager, which has been in flight since 1977; Galileo, which is ready to be launched; and CRAF, which is under design. These show the progression chronologically. Due to different mission require-

²J. M. Brayshaw, detailed weight tabulation computer printout (internal document), Jet Propulsion Laboratory, Pasadena, California, August 18, 1977.

³R. M. Dickinson, interoffice memorandum to E. C. Posner, IOM 860326 (internal document), Jet Propulsion Laboratory, Pasadena, California, March 31, 1986.

⁴G. W. Garrison, private communication, Jet Propulsion Laboratory, Pasadena, California, July 14, 1986.

ments, certain features of the spacecraft make direct comparison of percentage figures difficult. The most significant feature is that of the differences between the three propulsion systems. The propulsion subsystem of the Voyager mission module is only 35.26 kg. The main provider of propulsion, the propulsion module (158.55 kg), is jettisoned en route. Galileo, in contrast, has a very large retro-propulsion module (201.52 kg) included in the mission module.^{5,6} CRAF has an even heavier retro-propulsion module (374.73 kg) compared to Galileo.⁷ Other significant features are the spin-bearing assembly and probe-related hardware on board Galileo.

⁵*Galileo*, Project Document 625-205, Functional Requirements Book, vols. 1 and 2 (internal document), Jet Propulsion Laboratory, Pasadena, California, 1982.

⁶*Galileo Quarterly Mass Report and Equipment List*, issue 31 (internal document), Jet Propulsion Laboratory, Pasadena, California, April 29, 1986.

⁷*Mariner Mark II Configuration, Mass and Power Report*, issue 12 (internal document), Jet Propulsion Laboratory, Pasadena, California, August 18, 1986.

However, the present analysis approach is consistent throughout. It is felt that estimates of spacecraft mass are all within 5 percent error. In any case, the results give a fair comparison of mass among the three spacecraft studied.

The results show that the portion of the spacecraft attributable to telecommunications is substantial. In particular, the mass fraction for the three chronologically representative spacecraft, Voyager, Galileo, and CRAF, is 34 percent, 19 percent, and 18 percent, respectively. The large reduction of telecommunications mass from Voyager to Galileo is mainly due to the use of a deployable antenna instead of the solid antenna on Voyager.

We conclude that we should work toward further reducing the spacecraft telecommunications mass. Alternately, we can improve the telecommunications capability of the Deep Space Network (DSN) so that the required spacecraft telecommunications system mass can be reduced for equivalent communications performance. Continued progress in the area of deep space telecommunications technology development is essential to achieving new goals in space exploration.

Table 1. Telecommunications mass calculation for STRU

Category	Mass, kg	Contribution, %	Mass, %
Fully telecommunication	63.16	100.00	63.16
Non-telecommunication	42.10	0.00	0.00
Partly telecommunication			
PWR	21.16	39.64	8.39
CCS	5.22	10.00	0.52
FDS	5.34	10.00	0.53
AACCS	11.30	5.00	0.57
PROP	15.38	33.71	5.19
DSS	5.24	100.00	5.24
Total	168.63 kg		83.60 kg

Table 2. Voyager mass results summary

Subsystem	Total mass, kg	TC-related mass, kg	Non-TC mass, kg	TC-related percentage
STRU	168.63	83.32	85.31	49.41
RFS	44.44	44.44	0.00	100.00
MDS	8.41	8.41	0.00	100.00
PWR	136.39	54.07	82.32	39.64
CCS	15.51	1.55	13.96	10.00
FDS	19.32	4.05	15.27	20.95
AACS	49.74	2.49	47.25	5.00
PYRO	5.34	0.00	5.34	0.00
CABL	51.62	8.58	43.04	16.62
PROP	35.26	11.89	23.37	33.71
TEMP	29.63	6.31	23.32	21.29
DEV	16.12	1.29	14.83	8.02
DSS	15.15	14.39	0.76	95.00
SXA	5.09	5.09	0.00	100.00
SCI	123.00	0.00	123.00	0.00
SAH	5.68	0.00	5.68	0.00
Total	729.33	245.88	483.45	33.71

Table 3. Acronyms and abbreviations

Voyager Mission Module Subsystems:

RFS	Radio Frequency Subsystem
MDS	Modulation/Demodulation Subsystem
SXA	S/X-Band Antenna Subsystem
CCS	Computer Command Subsystem
FDS	Flight Data Subsystem
AACS	Attitude and Articulation Control Subsystem
STRU	Structure Subsystem
CABL	Cabling Subsystem
TEMP	Temperature Control Subsystem
PWR	Power Subsystem
PROP	Propulsion Subsystem
DSS	Data Storage Subsystem
PYRO	Pyrotechnic Subsystem
DEV	Mechanical Devices Subsystem
SAH	Systems Assembly Hardware
SCI	
CRS	Cosmic Ray Subsystem
PRA	Planetary Radio Astronomy Subsystem
PWS	Plasma Wave Subsystem
LECP	Low Energy Charged Particle Subsystem
PPS	Photopolarimeter Subsystem
PLA	Plasma Subsystem
UVS	Ultraviolet Spectrometer Subsystem
MAG	Magnetometer Subsystem
ISS	Image Science Subsystem
IRIS	Infrared Interferometer Spectrometer and Radiometer Subsystem

Miscellaneous Acronyms and Abbreviations:

CRAF	Comet Rendezvous Asteroid Flyby
TC	Telecommunications
DSN	Deep Space Network

Table 4. Voyager mass calculation

Subsystem		TC-related 100.0%	Non-TC 0.0%	PWR* 39.64%	CCS 10.0%	FDS 10.0%	AACS 5.0%	PROP 33.71%	DSS 100.0%	Subsystem total	TC-related total
STRU	Total	63.16	42.10	21.16	5.22	5.34	11.30	15.38	5.24	168.63	
	TC related	63.16	0	8.39	0.52	0.53	0.57	5.19	5.24		83.56
RFS	Total	44.44	0	0	0	0	0	0	0	44.44	
	TC related	44.44	0	0	0	0	0	0	0		44.44
MDS	Total	8.41	0	0	0	0	0	0	0	8.41	
	TC related	8.41	0	0	0	0	0	0	0		8.41
PWR	Total	0	0	136.39	0	0	0	0	0	136.39	
	TC related	0	0	54.07	0	0	0	0	0		54.07
CCS	Total	0	0	0	15.51	0	0	0	0	15.51	
	TC related	0	0	0	1.55	0	0	0	0		1.55
FDS	Total	2.35	0	0	0	16.97	0	0	0	19.32	
	TC related	2.35	0	0	0	1.70	0	0	0		4.05
AACS	Total	0	0	0	0	0	49.74	0	0	49.74	
	TC related	0	0	0	0	0	2.49	0	0		2.49
PYRO	Total	0	5.34	0	0	0	0	0	0	5.34	
	TC related	0	0	0	0	0	0	0	0		0
CABL	Total	2.47	25.64	6.09	2.07	5.03	2.86	6.93	0.53	51.62	
	TC related	2.47	0	2.42	0.21	0.50	0.14	2.34	0.50		8.58
PROP	Total	0	0	0	0	0	0	35.26	0	35.26	
	TC related	0	0	0	0	0	0	11.89	0		11.89
TEMP	Total	2.48	13.78	2.55	1.09	1.09	4.87	2.00	1.77	29.63	
	TC related	2.48	0	1.01	0.11	0.11	0.24	0.68	1.68		6.31
DEV	Total	0	12.86	3.26	0	0	0	0	0	16.12	
	TC related	0	0	1.29	0	0	0	0	0		1.29
DSS	Total	0	0	0	0	0	0	0	15.15	15.15	
	TC related	0	0	0	0	0	0	0	14.39		14.39
SXA	Total	5.09	0	0	0	0	0	0	0	5.09	
	TC related	5.09	0	0	0	0	0	0	0		5.09
SCI	Total	0	123.00	0	0	0	0	0	0	123.00	
	TC related	0	0	0	0	0	0	0	0		0
SAH	Total	0	5.68	0	0	0	0	0	0	5.68	
	TC related	0	0	0	0	0	0	0	0		0
Total										729.33 kg	246.14 kg

*Also see table in Appendix A.

Appendix A

Spacecraft Telecommunications System Power Estimates

Power is also an important limiting parameter for present-day planetary spacecraft design. A study of the power attributed to the communications system for spacecraft designs used in recent missions is presented here.⁸⁻¹¹

The basic approach is similar to that of the mass study, i.e., to examine the spacecraft by subsystem and allocate a portion of each subsystem to telecommunications. The percentage for power is calculated using power allocations derived from actual preflight subsystem testing (except in the case of CRAF, which is done using estimates). Only the dry mission module is taken into account. The portion of power for Voyager, Galileo, and CRAF attributable to the telecommunications system is 40 percent, 29 percent, and 18 percent, respectively. The results show that the portion of the spacecraft power attributable to telecommunications is substantial.

The task at hand is to separate the spacecraft into two parts, telecommunications and non-telecommunications, with regard to power. Again, the Voyager spacecraft is used to illustrate the analysis. This method is used to obtain results for the other spacecraft.

The telecommunication-related power percentage is calculated using figures from volume 1 of Project Document 618-205.⁸ These list 50 power modes, from launch through the Saturn encounter, with power allocations by subassembly

⁸*Mariner Jupiter/Saturn 1977*, Project Document 618-205, vol. 1 (internal document), Jet Propulsion Laboratory, Pasadena, California, pp. 11-23, June 8, 1977.

⁹*Mariner Jupiter/Saturn 1977*, Project Document 618-205, vol. 2 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 12, 1977.

¹⁰*Galileo Quarterly Power Report*, issue 33 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 20, 1986.

¹¹*Comet Rendezvous Asteroid Flyby*, Project Document 699-100, Rev. C (JPL D-1457 Rev. B) (internal document), Jet Propulsion Laboratory, Pasadena, California, May 1986.

and subsystem. Of the 50 modes, nine are identified as the main power modes, i.e., essentially the cruise background modes and any modes lasting more than two days. The figures from these nine modes are used for the Voyager power calculation and are summarized in Table A-1. Included in Table A-2 are the equations used to calculate the percentage of telecommunication-related power from each of the nine modes. The percentage of telecommunication-related power is extracted from each mode in a manner similar to the mass calculation. The telecommunication percentage for each mode is multiplied by the length of its respective mode and then summed. This sum is divided by the sum of the lengths of the modes to give an overall percentage.

The power figures from volume 1 of Project Document 618-205⁸ represent the maximum steady-state power allocated to spacecraft subsystems.

The lengths of the modes are derived from volumes 1 and 2 of Project Document 618-205^{8,9} using launch, Jupiter encounter, and Saturn encounter dates from the Voyager 2 mission (see Table A-1).

The power-needed calculation uses the Total DC Bus Power, which is the amount of power the subsystems will be using. The other percentage given is the power available using the net power capability (Most Probable) instead of Total DC Bus Power. Power available is slightly higher than power needed to provide a safe power margin. Therefore this gives a smaller overall percentage.

The mass analysis is done using both the power-needed and the power-available figures. The resulting telecommunications mass percentages are within 1 percent of each other, so only the power-needed figure was used in the preceding mass calculation.

Despite differences in the spacecraft, this analysis approach is consistent throughout. The present estimates of spacecraft power are all felt to be within 5 percent error. In any case, the results give a fair comparison of power among the three spacecraft studied.

Table A-1. Voyager power results summary

Mode number	Length, days	TC percentage of power	
		Power needed	Power available
24	96	46.97	45.31
25	541	38.70	33.20
27	6	39.31	35.56
31	40	45.28	47.32
36	39	45.75	44.73
37	700	38.62	34.44
39	8	33.90	29.35
43	29	45.28	49.50
48	29	37.69	37.89
Total		39.64	35.64

Table A-2. Power calculation equations

Category	Percentage
TC	100
PWR	39
CCS	10
FDS	10
AACS	5
PYRO	0
DSS	95
SCI	0
PROP	33

$$A = (MDS \times TC\%) + (PWR \times PWR\%) + (CCS \times CCS\%) + (FDS \times FDS\%) + (AACS \times AACS\%) + (PYRO \times PYRO\%) + (DSS \times DSS\%) + (STRU^*) + (PROP \times PROP\%) + (SCI \times SCI\%)$$

where $STRU^* = (\text{Bay 1 heater} \times TC\%) + (\text{Bay 2 heater} \times DSS\%) + (\text{Bay 6 heater} \times AACS\%)$

$$B = \text{AC wiring loss} \times \frac{A}{\text{total eng load} + \text{SCI}}$$

$$C = (2.4 \text{ kHz inv loss} \dots + \text{power factor loss}) \times \frac{(A + B)}{\text{total 2.4 kHz inv load}}$$

Total TC AC power load = A + B + C

$$D = (RFS \times TC\%) + (PWR \times PWR\%) + (FDS \times FDS\%) + (AACS \times AACS\%) + (SCI \times SCI\%) + (TC/S \times SCI\%) + (TC/E^*)$$

where $TC/E^* = (\text{Az Act Rep} \times SCI\%) + (\text{Sun Sen Htr} \times AACS\%) + (\text{IPU Valve Htr} \times PROP\%) + (\text{IPU Red V Htr} \times PROP\%) + (\text{IPU Thrus Htr} \times PROP\%) + (\text{TCAPU Red Htr} \times PROP\%) + (\text{Scn Pltfm Htr} \times SCI\%)$

$$E = \text{DC wiring loss} \times \frac{D}{\text{total reg DC load}}$$

Total TC DC power load = D + E

$$\text{TC power \%} = \frac{\text{total TC AC power load} + \text{total TC DC power load}}{\text{total DC bus power}}$$

Phase Calibration Generator

E. H. Sigman

Tracking Systems and Applications Section

A phase calibration system has been developed for the Deep Space Stations to generate reference microwave comb tones which are mixed in with signals received by the antenna. These reference tones are used to remove drifts of the station's receiving system from the detected data. This phase calibration system includes a cable stabilizer which transfers a 20 MHz reference signal from the control room to the antenna cone. The cable stabilizer compensates for delay changes in the long cable which connects its control room subassembly to its antenna cone subassembly in such a way that the 20 MHz is transferred to the cone with no significant degradation of the hydrogen maser atomic clock stability. The 20 MHz reference is used by the comb generator and is also available for use as a reference for receiver LOs in the cone.

I. Introduction

A Phase Calibration Generator (PCG) has been developed for the Deep Space Stations to provide phase calibration of the station's receiving system. This subsystem generates high stability microwave comb tones in the antenna cone which are referenced to the hydrogen maser in the control room. These comb tones are mixed with the signal received by the antenna. When the data from the received signal is processed, these tones are extracted from the data stream and their phase is used to determine the phase error which has been added to the data stream by the station's receiving system.

The PCG consists of a Transmitter unit in the control room and a Receiver unit in the antenna cone area. The Receiver unit contains the comb generator which generates tones at every integer multiple of $5/N$ MHz where N is an integer from 5 to 99. The PCG also contains a cable stabilizer which is split between the transmitting and receiving units. This cable stabilizer compensates for phase changes in the long cable connecting the Transmitter unit to the Receiver unit. The output of

the cable stabilizer consists of stabilized signals at $5/N$ MHz and at 20 MHz. The $5/N$ MHz is used by the comb generator to generate the comb tones. The stabilized 20 MHz is used as a frequency reference in the cone for generation of receiver LOs. A 5 to 20 MHz converter in the Transmitter unit provides the 20 MHz for the cable stabilizer. The $5/N$ MHz is supplied by a divide-by- N circuit in the Transmitter which is synchronized to the station's 1 pps time reference. This synchronization assures that the delay from the station's 1 pps to the comb generator pulse will always be the same constant value. A detailed description of the PCG design and performance follows. First the cable stabilizer will be discussed, and then the comb generator will be described.

II. Cable Stabilizer

A. Description

A block diagram of the cable stabilizer is presented in Fig. 1. More detailed block diagrams of the transmitting and receiving

units are shown in Figs. 2 and 3. The Transmitter generates 20 MHz and $5/N$ MHz signals from the 5 MHz station reference. These 20 and $5/N$ MHz signals are sent up the cable to the Receiver where they are separated by filters and mixed together in a double balanced mixer. The resultant $20 + 5/N$ MHz and $20 - 5/N$ MHz components are sent back down the cable to the Transmitter for phase comparison with the 20 MHz signal being sent up the cable. The error signal from this phase comparison is used to drive a voltage controlled phase shifter (VCPS) which compensates for phase changes in the cable. The result of this process is stabilized 20 MHz at the Receiver unit. The $5/N$ MHz into the Receiver unit is not cable stabilized but is used by the GATE circuit to gate individual cycles of the stabilized 20 MHz in order to obtain stabilized $5/N$ MHz pulses to drive the comb generator.

The power divider through which the $20 + 5/N$ MHz and $20 - 5/N$ MHz components are sent down the cable does not have perfect isolation. Thus some $20 + 5/N$ MHz and $20 - 5/N$ MHz components will appear on the cable-stabilized 20 MHz in the Receiver. Since the $5/N$ MHz is not cable stabilized, these components could cause phase instabilities in the $5/N$ MHz output of the GATE circuit. Therefore, before the cable-stabilized 20 MHz is sent to the GATE, it is filtered by a crystal filter which attenuates the $20 + 5/N$ MHz and $20 - 5/N$ MHz components down to a level at which they will not significantly affect the phase stability of the GATE circuit's output. The 20 MHz also passes through a second crystal filter before being made available as an output for generation of the Receiver LOs. This second crystal filter is necessary because of the stringent limits on the maximum allowable $20 + 5/N$ and $20 - 5/N$ components for the 20 MHz frequency reference.

The phase comparison process in the Transmitter requires a $5/N$ MHz signal which has approximately the same time delay with respect to the Receiver as do the $20 + 5/N$ and $20 - 5/N$ signals. This is provided by the Receiver, which has a high input impedance at $5/N$ MHz. Since this input impedance does not match the 50 ohm line, the $5/N$ MHz signal is reflected back down the cable. A hybrid in the Transmitter separates this reflected $5/N$ MHz signal from the $5/N$ MHz being sent up the cable.

The operation of the cable stabilizer may be described in more detail as follows. The signal sent up the cable has the form:

$$\sin 2\pi 20t$$

where t is in μsec .

This is multiplied by a $5/N$ MHz signal in the Receiver's double balanced mixer so that the signal sent down the cable to the Transmitter is

$$\sin (2\pi 20t + \phi_1) \sin 2\pi \frac{5}{N} t$$

where ϕ_1 = two way cable delay.

The first mixer in the Transmitter multiplies this by 20 MHz to give

$$\sin 2\pi 20t \sin (2\pi 20t + \phi_1) \sin 2\pi \frac{5}{N} t$$

The product of the first two sinusoids generates a 40 MHz term and a dc term. The 3 MHz low-pass filter (LPF) filters out the higher frequency terms leaving a term of the form

$$\cos \phi_1 \sin 2\pi \frac{5}{N} t$$

The second mixer in the Transmitter multiplies this by the reflected $5/N$ MHz to give

$$\cos \phi_1 \sin 2\pi \frac{5}{N} t \sin \left(2\pi \frac{5}{N} t + \delta \right)$$

where δ = error between the phase of the reflected $5/N$ MHz and the phase of the $5/N$ MHz modulation on the 20 MHz, and -45 degrees $< \delta < 45$ degrees.

The $5/N$ MHz phase error, δ , is due to the cable dispersion and hybrid imperfections. The product of the sinusoids generates a $10/N$ MHz term and a dc term. After low-pass filtering in the loop filter, the remaining term is

$$\cos \phi_1 \cos \delta$$

This signal is integrated by the loop filter to generate a control signal for the VCPS. The VCPS controls ϕ_1 thus completing the feedback loop. The action of the feedback loop is to keep

$$\cos \phi_1 = 0 \quad (1)$$

or

$$\phi_1 = \frac{\pi}{2} + 2n\pi \quad (2)$$

where n is an integer. As a result, the total delay through the VCPS plus the cable is held constant.

Note that the $5/N$ MHz phase error, δ , only affects the amplitude of the detected signal and does not affect the phase

at which lockup occurs as long as δ does not get close to ± 90 degrees. The use of the reflected $5/N$ MHz for demodulation assures that δ stays close to zero.

It can be seen that the feedback also has a potential lockup point for

$$\phi_1 = \frac{\pi}{2} + (2n + 1)\pi \quad (3)$$

However, this is an unstable lockup point because the feedback is positive. The type of feedback is determined by the slope of $\cos \phi_1 \cos \delta$. Thus, for $0 < \phi_1 < \pi$ the feedback is negative, whereas for $\pi < \phi_1 < 2\pi$ the feedback is positive. In the positive feedback region, the feedback loop will be unstable, which will cause the VCPS control voltage to swing to either its positive or its negative saturation limit. To avoid this problem, a search circuit is used. Whenever the VCPS control voltage gets too close to either its positive or its negative saturation limit, the search circuit takes over and sweeps the VCPS control voltage over its range until a stable lockup point is found. Since the VCPS is designed to have a range of slightly greater than 180 degrees (which makes its two way range slightly greater than 360 degrees), a stable lockup point will always be found. However, one must not get too close to the limits of the VCPS's range. If the VCPS control voltage gets too close to one of its limits, the search circuit will take over and find another lockup point near the other extreme of the VCPS's range. This will cause the output of the cable stabilizer to slip by one cycle. (The search circuit is described in more detail in a NASA technical brief which will be published soon.)

B. Error Sources

There are many error sources which can degrade the performance of the cable stabilizer. The most significant error sources which have been considered in the design of the cable stabilizer are discussed in the following sections.

1. Spurious signals. Spurious signals at any point in the cable stabilizer can alter the phase of the signal at that point in the circuit. The worst case is when the spurious signal is 90 degrees out of phase with the desired signal. Consider this case as shown in Fig. 4. The desired signal is the **D** vector and **S** is the spurious signal's vector. The resultant vector, **R**, is displaced in phase from the true value by the angle α . The phase error, α , is given by

$$\alpha = \tan^{-1} \left(\frac{S}{D} \right) \quad (4)$$

where **S** and **D** represent the vector magnitudes.

For **S** small with respect to **D**, this may be approximated by

$$\alpha = \frac{S}{D} \quad (5)$$

Let τ_α = the delay error.

Then, if f is the frequency of the desired signal,

$$\alpha = 2\pi f \tau_\alpha \quad (6)$$

so that

$$\tau_\alpha = \frac{S}{2\pi f D} \quad (7)$$

The design goal has been to keep the cumulative effect of all such errors less than 1 ps. From the above equation, any single such error will cause less than a 1 ps error at 20 MHz if $(S/D) < -78$ dB. Since there could easily be on the order of 10 such error sources in the cable stabilizer, if these error sources add in a root sum square manner, then each error source should have $(S/D) < -88$ dB. With this in mind, the cable stabilizer has been designed with the goal of keeping all spurious signals at least 90 dB down, and if possible 100 dB down.

Numerous spurious signal sources have been considered in the design of the cable stabilizer. For example, the 5 to 20 MHz converter in the Transmitter which generates the 20 MHz must have its spurious output components adequately suppressed. The 15 and 25 MHz components must be especially well suppressed since they are well within the bandpass of the 20 MHz bandpass filters (BPFs) of the cable stabilizer. Also, spurious signals from amplifiers in the 20 MHz paths of the cable stabilizer are kept within acceptable levels by using amplifiers with reasonably low distortion. These amplifiers are operated well below their maximum output level to assure that the harmonics they generate are kept down to an acceptable level.

In addition, interactions of components in the cable stabilizer have been considered. For example, spurious signals from one component which could generate unwanted intermodulation products in another component must be adequately suppressed. In particular, the signal voltages on the varicaps in the VCPS cause phase modulation of the signals in the VCPS. This generates intermodulation products of the 20 MHz and $20 + 5/N$ or $20 - 5/N$ MHz signals in the VCPS. These intermodulation products could generate intermodulation products in the mixer which might degrade the performance of the cable stabilizer. To avoid this problem, an extra 20 MHz BPF is used (shown on the left side of the VCPS in Fig. 2) to suppress intermodulation products generated by the VCPS. In

addition, the amplitude levels of the signals in the VCPS have been chosen to keep errors from intermodulation products down to acceptable levels.

2. Amplifier nonlinearities. An additional effect which can be caused by amplifier nonlinearities is a shift of the zero crossing of a signal. This is of concern since the GATE circuit in the Receiver converts the 20 MHz sine wave to a square wave by using a zero crossing detector. If the gain of an amplifier is asymmetric about zero volts, the amplifier will generate even harmonics and shift the zero crossing time of a sine wave. The amount of zero crossing shift will be dependent on the amplitude of the sine wave. If the harmonic distortion is low, the amount of zero crossing phase shift for a given amplitude change is, to a good approximation, proportional to the second harmonic distortion. Thus the zero crossing shift can be determined from the second harmonic distortion. The amplitude sensitivity of the zero crossing as a function of second harmonic distortion is given in Table 1 for a 20 MHz sine wave. Note that the phase error is independent of frequency. Thus the time error is inversely proportional to frequency.

To keep the zero crossing shift errors low, the 10 dB amplifier which drives the GATE has a 20 MHz BPF on its output to attenuate the amplifier harmonics. This filter keeps the second harmonic at least 75 dB down. From Table 1, the amplitude dependence of zero crossing time will then be kept down to 0.14 ps/dB. The main source of amplitude changes in the cable stabilizer is the VCPS, which has a loss which varies as its phase setting is changed. For normal operation, the loss variation is well under 1 dB so that zero crossing shifts due to amplitude changes are under 0.1 ps.

3. VSWR induced error

a. Analysis. Standing waves on the cable between the Transmitter and Receiver caused by a mismatch between the cable and the components connecting to the cable can reduce the accuracy with which the cable stabilizer can measure a change in cable length and correct for it. Consider the case shown in Fig. 5 where a signal with a voltage magnitude V is transmitted from the left end of the cable to the right end. A mismatch at the right end reflects the signal back down the cable with magnitude V' . The magnitude of the reflection coefficient is

$$\rho_2 = \frac{V'}{V} \quad (8)$$

Similarly, at the left end the signal V' is re-reflected back up the cable with a magnitude of V'' . The magnitude of the reflection coefficient at the left end is

$$\rho_1 = \frac{V''}{V'} \quad (9)$$

The signal received at the right end of the cable is the vector sum of V and V''

When the cable changes in length, the magnitude of the change can be determined by measuring the phase change of the signal received at the right end of the cable. The presence of the re-reflected signal, V'' , will corrupt this measurement. The error will be largest when V'' is in phase with V or 180 degrees out of phase with V . Consider the case where V'' is initially in phase with V at the right end of the cable. Then let the cable stretch slightly so that the signal V at the right end of the cable increases in phase by $\Delta\phi$. We wish to measure this phase change $\Delta\phi$ to determine the change in cable length and correct for it. Let us calculate the error in measuring $\Delta\phi$ caused by the re-reflected signal V'' . The signal V'' has traversed the cable three times by the time it reaches the right end so it will increase in phase by $3\Delta\phi$. Thus, after the cable stretch, the phase of V'' with respect to V will be $2\Delta\phi$. The phase relationship between these vectors is depicted in Fig. 6. The vector V_R , which is the vector sum of V and V'' , is the signal which will be detected at the right end of the cable. The phase angle, ϵ , between V_R and V is the error in measuring the phase of V and is thus the error in measuring $\Delta\phi$ from which the change in cable length is determined. From Fig. 6, the error angle ϵ is given by

$$\epsilon = \tan^{-1} \left(\frac{V'' \sin 2\Delta\phi}{V + V'' \cos 2\Delta\phi} \right) \quad (10)$$

From Eqs. (8) and (9)

$$V'' = \rho_1 \rho_2 V \quad (11)$$

so

$$\epsilon = \tan^{-1} \left(\frac{\rho_1 \rho_2 \sin 2\Delta\phi}{1 + \rho_1 \rho_2 \cos 2\Delta\phi} \right) \quad (12)$$

For small changes in cable length,

$$2\Delta\phi \ll 1$$

so

$$\sin 2\Delta\phi \approx 2\Delta\phi$$

$$\cos 2\Delta\phi \approx 1$$

We then get

$$\epsilon = \tan^{-1} \left(\frac{2\rho_1\rho_2\Delta\phi}{1 + \rho_1\rho_2} \right) \quad (13)$$

In the typical case

$$\rho_1\rho_2 \ll 1$$

so the above expression becomes

$$\epsilon = 2\rho_1\rho_2\Delta\phi \quad (14)$$

Let

$\Delta\tau$ = the change in cable time delay

τ_ϵ = error in measuring $\Delta\tau$

f = frequency of signal in cable

then

$$\Delta\phi = f\Delta\tau \quad (15)$$

$$\epsilon = f\tau_\epsilon \quad (16)$$

If we substitute Eqs. (15) and (16) into Eq. (14) we get

$$\tau_\epsilon = 2\rho_1\rho_2\Delta\tau \quad (17)$$

It is seen that the error in time delay measurement caused by a mismatch is independent of frequency.

Consider the case where

$$\rho_1 = \rho_2 = \rho \quad (18)$$

The voltage standing wave ratio (VSWR), S , is given by

$$S = \frac{1 + \rho}{1 - \rho} \quad (19)$$

so that

$$\rho = \frac{S - 1}{S + 1} \quad (20)$$

From the above expressions we can determine the worst case errors for measurement of cable delay change as a function of the VSWR. This worst case error is tabulated in Table 2 for a cable length change of 1 ns. In the Deep Space Stations the maximum cable length change over a 24 hour period is nor-

mally less than 1 ns so this table provides upper limits for VSWR induced errors.

b. Design. From the foregoing analysis it is seen that it is important to keep the VSWR of the cable from the Transmitter to the Receiver as low as possible along with all components which connect to the cable. Thus the VSWR of every component which connects to the cable has been carefully considered in designing the cable stabilizer. The 3 MHz Low-Pass Filters (LPFs) have tuned traps which isolate their circuitry from the cable at 20 MHz. If no other mismatches exist, the 3 MHz LPFs will degrade the VSWR to no worse than 1.02.

The 20 MHz BPFs use a circuit configuration which minimizes the VSWR at the center frequency and keeps the VSWR low in a symmetric manner about the center frequency. The physical layout of the circuit has been carefully determined for minimum VSWR, and each filter is individually tuned for minimum VSWR. Typical performance for these filters is a VSWR of <1.02 for 19.5 to 20.5 MHz and a VSWR of <1.06 for 19.0 to 21.0 MHz.

The VCPS has also been designed for low VSWR. The typical VCPS has a VSWR which drops to a minimum value of about 1.02 somewhere near the center of its phase correcting range. At the edges of its range the VSWR is <1.22. For each VCPS the control voltage at which it has minimum VSWR is determined in the lab. When the PCG is installed in the field, the length of the cable from the Transmitter to the Receiver is adjusted so that the cable stabilizer is operating with a control voltage close to this value. As a result of this procedure, the VCPS is normally in a range where its VSWR is <1.04. Then for the overall Transmitter the VSWR is normally <1.09.

For the Receiver, the input VSWR is normally <1.19. Using this number for the Receiver and 1.09 for the Transmitter, the reflection coefficients for the Transmitter and Receiver are:

$$\rho_T < 0.043$$

$$\rho_R < 0.087$$

If we put these numbers into Eq. (17) we get

$$\frac{\tau_\epsilon}{\Delta\tau} < 7.5 \text{ ps/ns}$$

for the worst case VSWR induced error.

The reciprocal of this number gives:

cable correction factor > 134

This number indicates the amount by which the cable stabilizer improves the cable performance. Both laboratory and field tests of the cable stabilizer have indicated that the typical cable correction factor obtained is 100 to 1000. The close agreement between measured performance and calculated performance seems to indicate that VSWR is the main limiting error source.

4. Dispersion. Dispersion in the cable or cable stabilizer components can generate errors in the cable compensation. The cable stabilizer feedback loop holds the two way delay in the cable constant. If the delay through the cable in each direction is equal, then the one way delay will be held constant. However, a typical cable will have some dispersion so that the delay at $20 + 5/N$ MHz and $20 - 5/N$ MHz will not be the same as the 20 MHz delay. But the $20 + 5/N$ and $20 - 5/N$ delays will differ from the 20 MHz delay by about the same amount except in opposite directions. Since the cable stabilizer uses both the $20 + 5/N$ and $20 - 5/N$ signals, the dispersion errors will almost cancel if the cable stabilizer has equal gain at these two frequencies and symmetric phase response about 20 MHz. The components of the cable stabilizer which limit its ability to meet these criteria are the 20 MHz BPF and the VCPS.

As mentioned previously, the 20 MHz BPF uses a circuit configuration which gives virtually symmetric behavior about the center frequency. The flatness of this filter is ± 0.075 dB from 19 to 21 MHz. The attenuation at $20 + 5/N$ MHz matches the attenuation at $20 - 5/N$ MHz within 0.1 dB. The VCPS is flat to better than ± 0.1 dB from 19 to 21 MHz. These performance specifications keep the gain of the cable stabilizer at $20 + 5/N$ MHz within 4 percent of the gain at $20 - 5/N$ MHz, thus minimizing dispersion effects.

5. Choice of cable stabilizer frequency. Since 5 MHz is the reference frequency used for the cable stabilizer, an integer multiple of 5 MHz is the logical choice for the cable stabilizer frequency. For a good crystal filter 20 MHz is about the highest practical frequency, so this was one of the driving factors in the choice of frequency. In addition, 20 MHz is a low enough frequency for the cable stabilizer to handle any expected cable length changes without a cycle slip. In practice, for typical cable length changes the VCPS stays well within its range of optimal VSWR performance when 20 MHz is used.

The effect of cable stabilizer frequency on the cable stabilizer's ability to correct changes in cable length is also of interest. As indicated in previous sections, errors due to spurious signals and amplifier nonlinearities decrease with increas-

ing frequency. Also, phase noise tends to be constant with frequency so the time jitter caused by phase noise will be lower for higher cable stabilizer frequencies. However, the VSWR induced error is a function only of VSWR and is independent of frequency. Since the VSWR of most electronic circuits tends to increase with frequency, the VSWR induced error tends to increase with frequency. As indicated previously, in this cable stabilizer the VSWR induced error appears to be the dominant error source. Thus, if cable correction ability is the primary criterion used to select the cable stabilizer frequency, one would probably not want to choose a higher frequency than 20 MHz. In fact, it may be possible to improve the cable correction ability by a factor of 2 or so by going to a lower frequency. However, use of a lower frequency would increase phase noise effects. Thus, one would have to weigh phase noise effects against cable correction ability if another operating frequency for the cable stabilizer were to be considered.

6. Thermal drifts. The long term stability of the PCG is primarily limited by thermal drifts of the electronic components. To minimize the effects of environmental temperature changes, all critical electronic circuitry is housed in ovens which maintain a relatively constant temperature.

a. Ovens. All of the ovens used in the PCG have a copper or aluminum baseplate which acts as a temperature controlled surface. All electronic modules are built in aluminum boxes which are mounted on the baseplate with thermal grease to assure good thermal conduction. A thermistor at the center of the baseplate senses the baseplate temperature and connects to a proportional controller which controls the oven temperature. An aluminum box encloses the electronic modules and the baseplate. This box is surrounded by a layer of insulation and then by an outer box.

The Transmitter oven has its inner box completely covered with pad heaters which provide uniform heat. A layer of air between the inner and outer boxes provides insulation. A fan circulates air around the outer box to remove excess heat. The oven holds a nominal internal temperature of 50°C for external temperatures of 0°C to 25°C . Any external temperature change is reduced by a factor of about 100 on the inside of the oven.

The Receiver oven uses thermoelectric heat pumps to control the interior temperature by pumping heat into or out of the baseplate depending upon the outside temperature. The aluminum baseplate is mounted on six copper rods. Each copper rod connects to a heat dissipator plate on the outside box through a thermoelectric heat pump. The space between the inner and outer boxes is filled with urethane foam and Styrofoam insulation. The oven maintains a nominal inside tempera-

ture of 60°C for outside temperatures in the range of -55°C to +55°C. External temperature changes are reduced by a factor of about 100 on the inside of the oven.

In some installations the Receiver is too large to be mounted close to the comb generator's microwave injection point on the antenna assembly. In such cases, the comb generator is mounted in its own small oven called a CGA (Comb Generator Assembly). The CGA also uses thermoelectric heat pumps. Its copper baseplate is mounted on four thermoelectric heat pumps which in turn are mounted on the external heat dissipation plate. Styrofoam insulation is used between the inner and outer boxes. The oven holds a nominal internal temperature of 60°C for external temperatures of -55°C to +65°C. The oven reduces external temperature changes by a factor of about 150.

b. Temperature coefficients of electronics. The temperature coefficients of the various modules in the PCG have been measured by changing the set point of the oven and measuring the resultant phase change. The phase change was determined by using the long term stability test setup described in Section II.C.2.b. The comb generator phase change was measured using the comb generator test setup described in Section III.C. For the LEVELING-AMP-GATE the temperature coefficient was not measured but was determined from known characteristics of the chips in the circuit. The typical temperature coefficients of delay for the major modules in the PCG are shown in Table 3.

The overall typical temperature coefficients for the Transmitter and Receiver electronics are:

Transmitter electronics	13 ps/°C
Receiver electronics	
To comb generator output	48 ps/°C
To 20 MHz output	80 ps/°C

The ovens reduce the sensitivity of the electronics to the environment. For each of the PCG components the typical sensitivity to the environment is:

Transmitter	0.6 ps/°C
Receiver	
To comb generator output	0.5 ps/°C
To 20 MHz output	0.8 ps/°C
CGA	0.01 ps/°C

Note that the Transmitter oven reduces the environmental sensitivity of the electronics to about 0.1 ps/°C. However, the

cable in the Transmitter chassis which brings the 5 MHz input from the back panel to the oven is not a phase stable cable and is not protected from the environment. This cable will add about 0.5 ps/°C to the environmental sensitivity of the Transmitter.

C. Performance

1. Cable correction ability. The cable correction ability of the cable stabilizer has been tested in the laboratory by using the long term stability test setup which is described in Section II.C.2.b. A piece of cable with a known delay of about 2 ns was added to the cable between the Transmitter and Receiver, and the delay change in the Receiver's output was measured. This test has been performed for various values of N as part of the acceptance testing for each unit which has been installed in the field. The results of these tests indicate that the normal range of cable correction ability is 100 to 1000.

The first two PCGs which were installed at the Goldstone Deep Space Station were also tested in the field for cable correction ability. The phase of the phase calibrator tone was monitored through the station's receiving system while a cable with a 4.3 ns delay was added to the cable from the Transmitter to the Receiver. The results of this test are shown in Table 4. This test also showed the cable correction ability to be in the range of 100 to 1000.

2. Long term stability

a. Expected performance. The following estimates for the overall performance of the PCG subsystem in a Deep Space Station are based upon the data presented in the previous sections. The cable from the Transmitter to the Receiver is taken to be a 1000-foot-long hard line. Such a cable would typically have a temperature coefficient of 25 ppm/°C and a delay of 1.5 microseconds. This would give a temperature coefficient for the cable delay of 37.5 ps/°C. Since the cable stabilizer corrects the cable by a factor of at least 100, the corrected cable would have a temperature coefficient of less than 0.375 ps/°C.

A setup is considered in which the comb generator is separate from the Receiver in its own CGA oven. It is assumed that the CGA is connected to the Receiver through a 10-foot-long phase stable cable with a temperature coefficient of 14 ppm/°C. The delay through such a cable would be 15 ns, which would give a delay temperature coefficient of 0.2 ps/°C.

The coupling of the comb generator into the front end of the station's microwave receiving system is also considered. The present coupler is a loop coupler which does not have adequate performance for phase calibration applications. This coupler is going to be replaced with a Bethe hole coupler on the antenna's feed horn. Aside from the improved performance which will be obtained, an important advantage of the Bethe

hole coupler is that it will inject the phase calibrator tone as far forward in the receiving system as is possible.

The estimated environmental temperature changes for the PCG subsystem over a 24 hour period are taken as:

Transmitter	$\pm 1^{\circ}\text{C}$
Cable (from Transmitter to Receiver)	$\pm 5^{\circ}\text{C}$
Cone (includes Receiver, cable from Receiver to CGA, CGA, and Coupler)	$\pm 15^{\circ}\text{C} (\pm 5^{\circ}\text{C})$

For the cone the estimates indicated are for poor environmental control. Numbers in parentheses indicate performance for more reasonable control of cone temperature. With these temperature variations, the expected delay variations over 24 hours can be calculated from the temperature coefficients and are shown in Table 5.

It is seen that the main limitations on performance are probably the loop coupler and temperature control of the cone area. When the Bethe hole coupler is installed, the PCG should meet its specification of 10 ps for delay variations over 24 hours. Reasonable control of the cone temperature will assure that the PCG is well within this specification.

Over a 1000 second period the estimated environmental temperature changes for the PCG subsystem are taken as:

Transmitter	$\pm 0.5^{\circ}\text{C}$
Cable (from Transmitter to Receiver)	$\pm 0.16^{\circ}\text{C}$
Cone (includes Receiver, cable from Receiver to CGA, CGA, and Coupler)	$\pm 0.5^{\circ}\text{C} (\pm 0.16^{\circ}\text{C})$

For the cable and the cone, these estimates are derived by assuming that the variations shown above for 24 hours occur linearly over an 8 hour period. From these temperature variation estimates the expected delay variations over 1000 seconds can be calculated and are shown in Table 6. In Table 6 the calculated delay variations are divided by 1000 sec to obtain the expected Allan variance.

We can also obtain expected Allan variance for the 20 MHz output of the cable stabilizer which is used as a reference for receiver LOs. From the temperature variations shown above and the previously quoted temperature coefficients the results shown in Table 7 were obtained.

Note that the cone temperature variation is probably the primary limiting factor for cable stabilizer performance. If temperature variations in the Transmitter are significant, an improvement in performance may be attainable if the cable bringing the 5 MHz into the Transmitter oven is replaced with a shorter cable or a phase stable cable.

b. Measured performance. The long term stability of cable stabilizers was tested in the laboratory with the setup shown in Fig. 7. This setup measures the phase of the 20 MHz output with respect to the 5 MHz input. The delay resolution on the output is about 1 ps. As part of the acceptance testing for each PCG which has been built, the cable stabilizer is left running for at least 24 hours and its performance is recorded. The delay at different points in time is read off the strip chart and the Allan variance is calculated from these delay points. The typical Allan variance which has been seen for the 20 MHz output of the cable stabilizer is shown in Table 8.

Allan variance tests have also been performed on the cable stabilizer at the Maser Test Facility of the Frequency and Timing Subsystem Group. For these tests the output of a cable stabilizer running off one hydrogen maser was compared with another maser whose output was down-converted to 20 MHz. The results showed no measurable degradation of the maser stability. In these tests the Allan variance for each of the masers alone was 1×10^{-15} over 1000 sec.

The long term stability tests in our laboratory showed that the long term behavior was dominated by a drift which was almost linear with time and had a drift rate which decreased with time. In one typical case, the drift was 41 ps for the first 24 hours, 19.5 ps for the second 24 hours, and 17.5 ps for the third 24 hours. This drift is caused by the crystal filter. The crystal filter manufacturer does burn in the crystal filters for a few months to reduce aging effects. However, each time a crystal filter is brought up to operating temperature it will take some time to stabilize. No long term stability tests have been performed in our laboratory after more than a 3 day warm-up. However, in normal operation, the oven will be kept at operating temperature on a continuous basis. This will minimize crystal filter drifts.

3. Phase noise. The phase noise of the cable stabilizer was measured by mixing the output of two cable stabilizers together in a mixer. The phase of one cable stabilizer was delayed in order to zero out the dc component in the mixer output. The remaining output of the mixer, which is the phase noise, was amplified and measured. The typical total phase noise measured in a 3 MHz bandwidth was:

20 MHz output	0.4 ps
20 MHz to Leveling-Amp-Gate	0.6 ps

The phase noise spectrum was measured at the Maser Test Facility of the Frequency and Timing Subsystem Group. A similar procedure was followed except that a digital spectrum analyzer was used as the measuring device. The typical results that were obtained are shown in Table 9.

III. Comb Generator

A. Description

The microwave comb generator generates a comb spectrum from 2 to 10 GHz. The input is the $5/N$ MHz pulses from the cable stabilizer which have a repetition rate of up to 1 MHz. The comb generator uses a step recovery diode (SRD) which is pulsed at the input repetition rate to generate the comb tones. Whenever the diode current switches from forward bias to reverse bias, the diode continues conducting for a time period equal to τ , the minority carrier lifetime, and then very quickly switches to the nonconducting state in 50 ps. This rapid change in diode current induces a 50 ps voltage pulse in an inductor which is in series with the diode. This voltage pulse, which is the output of the comb generator, has harmonics up through 10 GHz which are coherent with one another in phase.

The time of occurrence of the comb generator's output pulse depends upon the τ of the SRD. The τ is dependent on temperature and on the ratio of forward SRD current to reverse SRD current. The forward and reverse currents of the SRD are set by a high stability dc bias circuit which precisely controls their ratio in order to keep τ at a value of about 1.5 ns. The temperature coefficient of τ is 1 percent/ $^{\circ}\text{C}$ so the temperature dependence of the SRD would be 15 ps/ $^{\circ}\text{C}$. A thermistor in the dc bias circuit changes the ratio of the forward and reverse currents in a manner which compensates for the SRD's temperature dependence at the operating temperature of 60°C in order to minimize temperature dependence. Tests of the comb generator have shown that the overall temperature coefficient at 60°C is about 1 ps/ $^{\circ}\text{C}$.

The magnitude of the SRD output can be set by a digital control word. The forward and reverse SRD currents are determined by the dc bias circuit which maintains them at values proportional to the digital control word. The ratio of the forward and reverse currents is always maintained at the same constant value except for temperature compensation.

B. Dispersion

The output of the comb generator must have very low dispersion, that is, very small deviation from a linear phase versus frequency characteristic. For VLBI (Very Long Baseline Interferometry) the performance goal for the comb gener-

ator is to limit dispersion to ± 0.88 degree over a 400 MHz bandwidth.

The main cause of dispersion is VSWR in the microwave circuitry. Consider the cable in Fig. 5 where a signal V is being sent from the left end to the right end. Any mismatch at the right end will reflect a signal V' back down the cable with a reflection coefficient of ρ_2 . Similarly, at the left end a signal V'' is reflected back up the cable with a reflection coefficient of ρ_1 . At the right end of the cable the re-reflected signal, V'' , will add vectorially to V and alter its phase by an amount ϵ . This situation is shown in Fig. 6 where the two way cable delay, $2\Delta\phi$, is the phase difference between V and V'' at the right end of the cable. Since the cable delay, $\Delta\phi$, is a function of frequency, the phase deviation of the signal, ϵ , will be a function of frequency. The maximum positive value of ϵ will occur when $2\Delta\phi = 90$ degrees, and the maximum negative value will occur when $2\Delta\phi = -90$ degrees. The frequency span over which ϵ changes from a maximum positive value to a maximum negative value will be the frequency for which the cable is one quarter wavelength long.

Since it is desirable to make the rate of change of ϵ as small as possible with respect to frequency, it is best to make the cable as short as possible. This will be the case when the Bethe hole couplers are installed in the Deep Space Stations. The CGA unit will be physically mounted right next to the Bethe hole coupler. Its output connector will connect directly to the coupler's connector with no intervening cable. The effective cable from the comb generator output to the coupler will then be about 15 cm long. The resultant frequency span from maximum positive ϵ to maximum negative ϵ will be about 300 MHz. Thus, ϵ will always have at least one maximum in a 400 MHz frequency span. In order to meet VLBI requirements, the maximum ϵ should be less than 0.88 degree. The maximum value of ϵ can be determined by putting $2\Delta\phi = 90$ degrees into Eq. (12) to get

$$\epsilon = \tan^{-1}(\rho_1\rho_2) \quad (21)$$

or

$$\rho_1\rho_2 = \tan \epsilon \quad (22)$$

For $\epsilon < 0.88$ degree

$$\rho_1\rho_2 < 0.0154$$

Consider

$$\rho_1 = \rho_2 = \rho$$

Then

$$\rho < 0.124$$

So from Eq. (19) the requirement on the VSWR is

$$S < 1.28$$

The coupler which injects the comb tones into the microwave receiving system must meet this requirement. The comb generator output must also meet this requirement. Since the SRD is a very poor match to a 50 ohm line, the comb generator has a 10 dB attenuator on its output to isolate the SRD from the line. The output VSWR of a 10 dB attenuator is always < 1.25 no matter how poor its input match. The line from the SRD to the 10 dB attenuator input has a poor match at the SRD end, but its length is kept very short (< 3 mm) so that the phase variation with frequency is minimized.

C. Performance

The long term stability of the comb generator has been measured with the setup shown in Fig. 8. A PCG Transmitter and Receiver were set up with a Comb Generator inside the Receiver. A second comb generator inside a separate oven was also connected to the Receiver's cable stabilizer output. A power switch operating off a 1 pulse per second (1 pps) signal alternated power between the two comb generators so that one comb generator operated on even seconds and the other comb generator operated on odd seconds. A power combiner connected the outputs of the two comb generators to a receiving setup consisting of a microwave receiver, IF Converter, and Formatter to down-convert the microwave signal and digitize it. The digital data stream went to a Digital Tone Extractor (DTE) which determined the phase of the phase calibration tone. The DTE was synchronized with the 1 pps and determined the phase every second by integrating over the last half of each second. This allowed the phase of each comb generator to be tracked separately over time.

Two tests were performed with the test setup. In the first test the temperature coefficient of the comb generator was determined by varying the set point of the second oven and observing the change in relative phase of the two comb gener-

ators. In the second test the oven temperatures were held constant and the phase behavior of the two comb generators was recorded over time. The Allan variance of the phase difference between the two comb generators was determined for various time intervals. This Allan variance was divided by the square root of 2 to determine the Allan variance for a single comb generator. The results of these tests are shown in Table 10. It can be seen that the stability of the comb generator is significantly better than the cable stabilizer. Thus, for the entire PCG the stability is limited by the cable stabilizer.

IV. Very Long Baseline Interferometry Results

In field usage the PCG has proven to be very useful and in some cases invaluable for improving the quality of VLBI data. Analysis of VLBI experiments in which the prototype PCG unit was used at Deep Space Station 13 has been performed by Chris Jacobs at JPL. For phase delay tests in the 8.4 GHz (X) band, he found that application of phase calibrator correction to the data reduced noise, thus making it easier to resolve ambiguities and connect phase points. In another experiment, the Traveling Wave Maser (TWM) was drifting badly and had to be retuned often. Application of the phase calibrator correction saved this experiment.

Some results of another experiment which was saved by the PCG are shown in Figs. 9 and 10. In this experiment the differential group delay between Deep Space Station 13 and 45 was being determined using 40 MHz Bandwidth Synthesis (BWS). The TWM at station 13 was erroneously set to 20 MHz bandwidth so that all the BWS channels being recorded were outside the TWM bandwidth. The differential group delay over time for the two outermost BWS channels is shown in Fig. 9. At each point the error bar due to system noise is shown. For the first part of the experiment it can be seen that the drift over time is large compared to the error bars, thus degrading the data. The same data after phase calibrator correction is shown in Fig. 10. It is seen that the drift is virtually gone. Without the phase calibrator correction there were many points for which the ratio of group delay error to system noise was on the order of 30. Phase calibrator correction dropped this ratio to less than 1, thus making the data usable.

Table 1. Amplifier non-linearity: $f = 20$ MHz

Second harmonic attenuation from fundamental (dB)	Zero crossing shift / Amplitude change (ps/dB)
40	8
50	2.5
60	0.8
70	0.25
75	0.14
80	0.08

Table 2. Cable change measurement error due to VSWR

VSWR	Error (ps)
1.05	1.18
1.10	4.52
1.15	9.74
1.20	16.5
1.25	24.7

Note: Worst case error in measurement of 1 ns change in cable delay.

Table 3. Temperature coefficients for modules

Module	Temperature coefficient (ps/°C)
5 to 20 MHz converter	1
10 dB amplifier	3
20 dB amplifier	6
10 dB amp-filter-coupler	25
20 dB amp-filter	29
1st crystal filter	10-14
2nd crystal filter	3-5
Leveling-amp-gate	2
Comb generator	1

Table 4. Field test of cable correction ability conducted at Goldstone Deep Space Station (cable delay increased by 4.3 ns)

Frequency band	N	Cable correction error (ps)	Correction ability
2.2 GHz (S)	5	4.4	977
2.2 GHz (S)	10	34.9	123
8.4 GHz (X)	5	11.9	361
8.4 GHz (X)	10	27.3	158

Table 5. Expected 24 hour stability for comb tone output

Component	Delay variation (ps)
Transmitter	±0.6
Stabilized cable	±1.9
Receiver	±7.5 (±2.5)
Cable to CGA	±3.0 (±1.0)
CGA	±0.15 (±0.05)
Coupler	Loop ±15.0 (±5.0) Bethe ±1.5 (±0.5)
RSS	17.2 ps (6.0 ps) 8.5 ps (3.4 ps)

Note: Numbers in parentheses indicate performance with moderate temperature control of cone.

Table 6. Expected 1000 second stability for comb tone output

Component	Delay variation (ps)			
Transmitter	±0.3			
Stabilized cable	±0.06			
Receiver	±0.25	(±0.08)		
Cable to CGA	±0.1	(±0.03)		
CGA	±0.005	(±0.002)		
Coupler	Loop ±0.5	(±0.16)		Bethe ±0.05 (±0.016)
RSS	0.65 ps (0.36 ps)		0.41 ps (0.32 ps)	
Expected Allan variance (1000 sec)	6.5 × 10 ⁻¹⁶	(3.6 × 10 ⁻¹⁶)		4.1 × 10 ⁻¹⁶ (3.2 × 10 ⁻¹⁶)

Note: Numbers in parentheses indicate performance with moderate temperature control of cone.

Table 7. Expected 1000 second stability for 20 MHz output

Component	Delay variation (ps)	
Transmitter	±0.3	
Stabilized cable	±0.06	
Receiver	±0.4	(±0.13)
RSS	0.50 ps	(0.33 ps)
Expected Allan variance (1000 sec)	5 × 10 ⁻¹⁶	(3.3 × 10 ⁻¹⁶)

Note: Numbers in parentheses indicate performance with moderate temperature control of cone.

Table 9. Measured phase noise spectrum: typical values for 20 MHz output

Frequency offset from carrier (Hz)	Phase noise in 1 Hz bandwidth (dB from carrier)
1	-107
10	-118
100	-125
1000	-125
10,000	-130

Table 8. Measured stability of 20 MHz output

Time interval (sec)	Allan variance
1	1.5 × 10 ⁻¹³
10	2 × 10 ⁻¹⁴
100	3 × 10 ⁻¹⁵
1000	4 × 10 ⁻¹⁶
10,000	1 × 10 ⁻¹⁶

Table 10. Measured comb generator stability

Temperature coefficient:	1 ps/°C	
Allan variance:	3 × 10 ⁻¹⁶	over 100 sec
	7 × 10 ⁻¹⁷	over 1000 sec
	1 × 10 ⁻¹⁷	over 10,000 sec

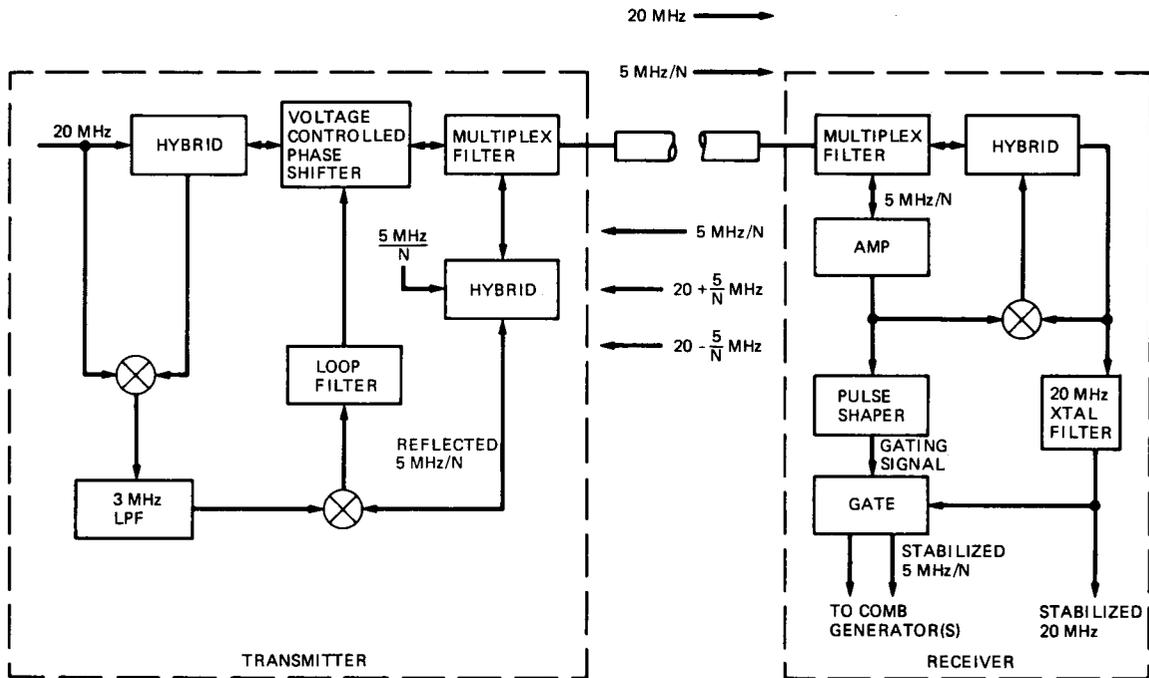


Fig. 1. Cable stabilizer

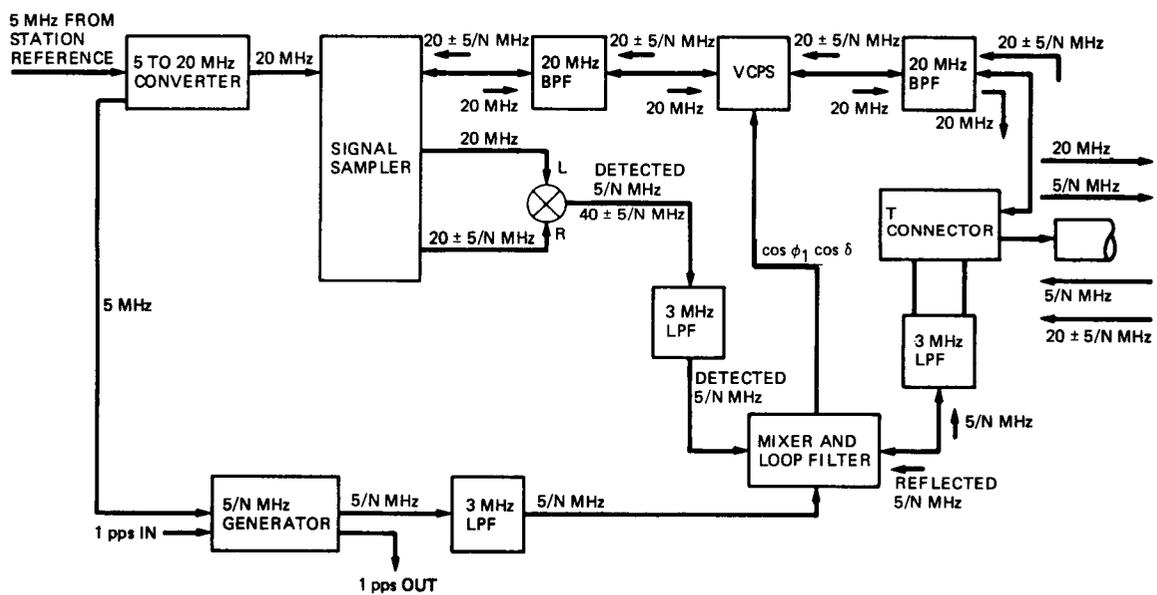


Fig. 2. PCG transmitter

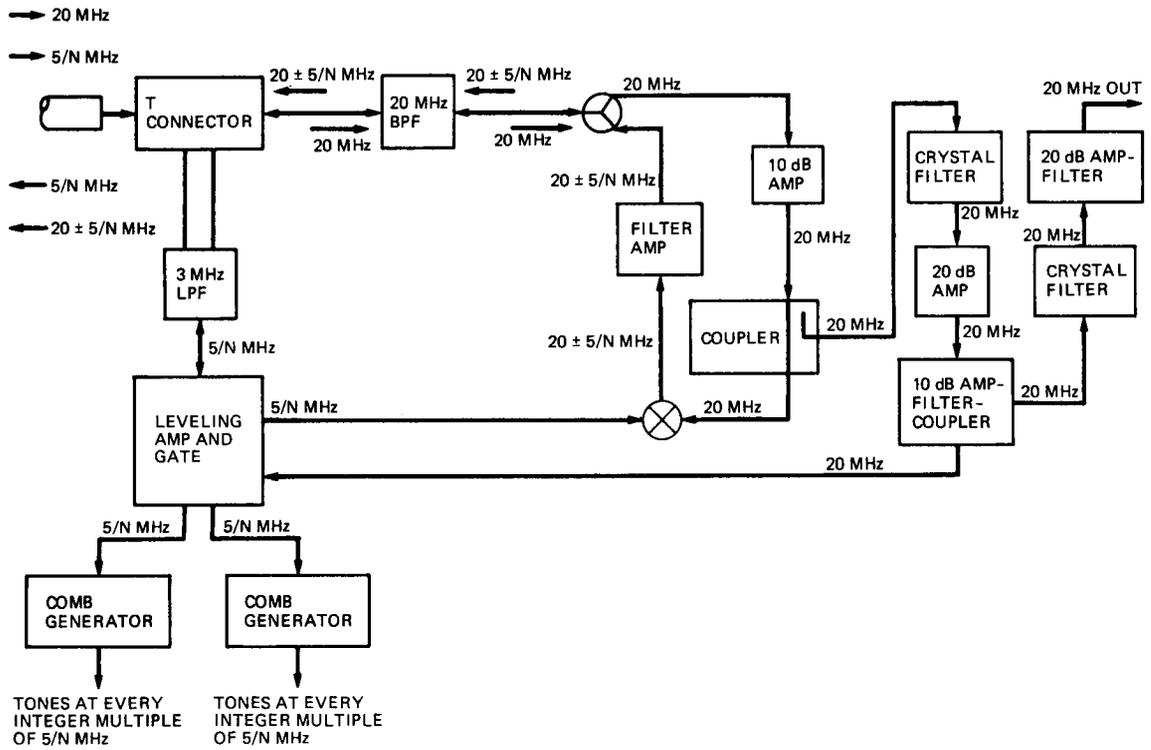


Fig. 3. PCG receiver

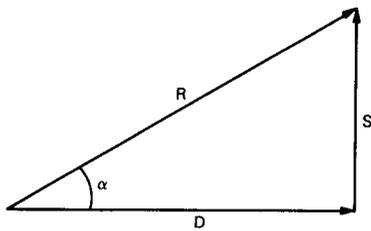


Fig. 4. Phase error due to spurious signals

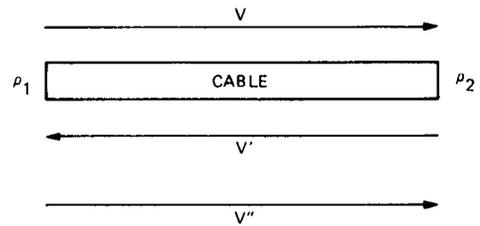


Fig. 5. Cable reflections

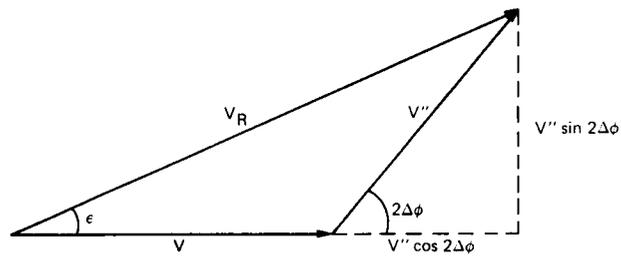


Fig. 6. Phase error due to cable reflections

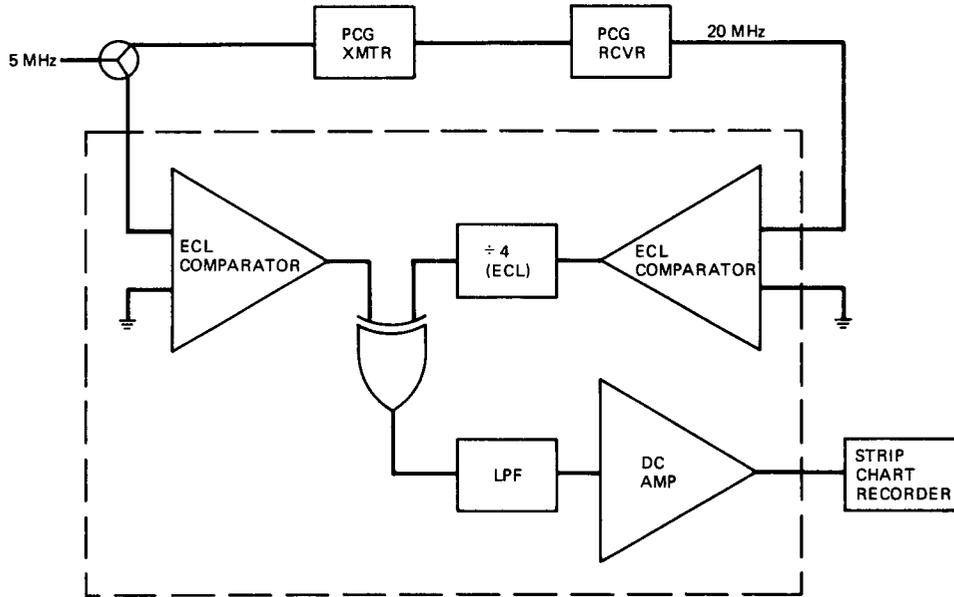


Fig. 7. Test setup for long term stability measurement

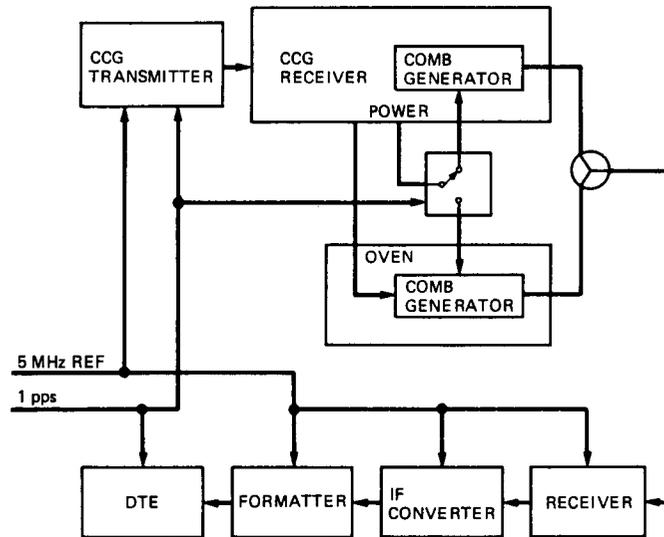


Fig. 8. Comb generator test setup

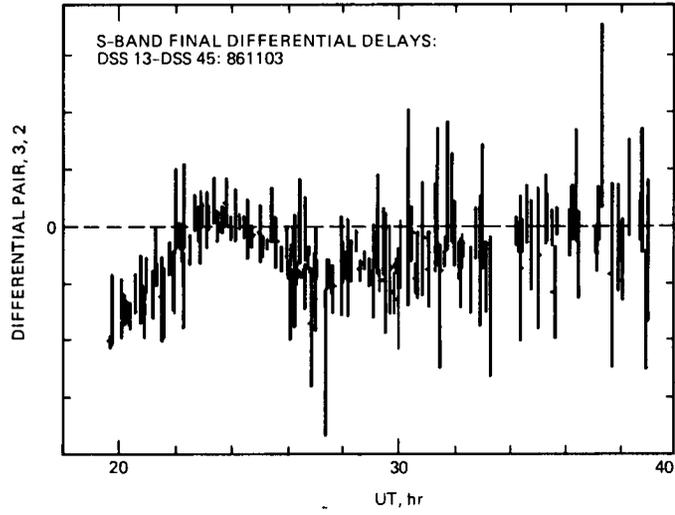


Fig. 9. VLBI results without phase calibration correction

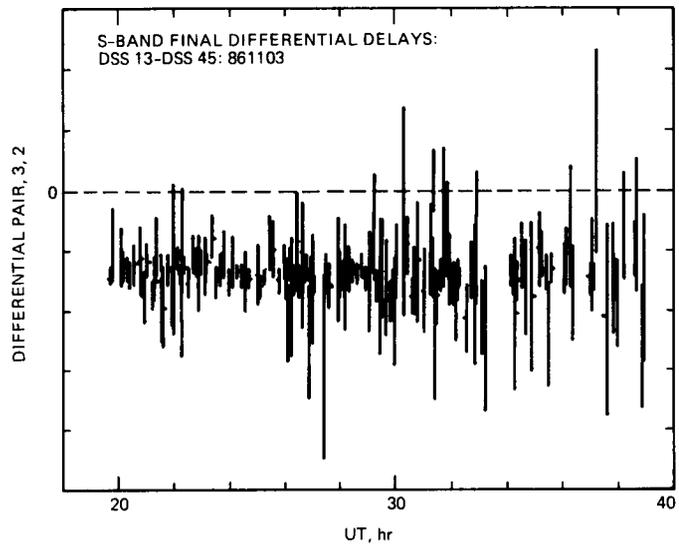


Fig. 10. VLBI results with phase calibration correction

A Simplified, General-Purpose Deep-Space Ranging Correlator Design

R. C. Tausworthe
Information Systems Division

J. R. Smith
DSN Data Systems Section

This article describes a much-simplified, yet more general-purpose multi-channel deep-space ranging system correlator design than has been used in past JPL spacecraft ranging systems. The method applies to detection of both single-component and multiple-component ranging codes, in either sequential (μ) or composite (τ) transmitted forms, and using either pseudonoise or square-wave components. Using this design, the Phobos Probe ranging system correlator computational complexity was reduced by over three orders of magnitude in multiply-and-add circuits and 45,000 bits of accumulator storage.

I. Introduction and Background

JPL Spacecraft ranging systems transmit, from the ground to the spacecraft, a periodic code $x(t)$ modulated on the uplink carrier. This signal received by the on-board transponder, is demodulated and retransmitted via the downlink to the ground station. The received version of the transmitted code at the ranging demodulator assembly interface appears attenuated by a factor α , delayed by the round-trip light-time τ (continuously changing due to the earth and spacecraft relative velocity), and immersed in noise, $n(t)$. Symbolically,

$$y(t) = \alpha x(t - \tau) + n(t) \quad (1)$$

Present DSN ranging systems [1] use the two-way detected Doppler signal from the receiver to program the clock oscilla-

tors of the receiver codes, generating replicas of the transmitted-code component(s) to maintain a constant value of τ between the received signal, $y(t)$, and the local receiver codes.

The optimum [2] estimate in the maximum-likelihood sense, $\hat{\tau}$, for the time delay τ over a given observation interval $[0, T]$, set by ranging accuracy requirements and the incoming signal-to-noise ratio,¹ is that value maximizing the cross-correlation integral

$$I_{\hat{\tau}} = \int_0^T y(t)c(t - \hat{\tau}) dt \quad (2)$$

¹Typically minutes to hours in duration.

where $c(t)$ denotes a particular code component of the transmitted signal, $x(t)$. Computation of $I_{\hat{\tau}}$ thus requires a ground receiver channel equivalent to the form shown in Fig. 1.

To determine $\hat{\tau}$, a set $\{\hat{\tau}_i : i = 0, \dots, N - 1\}$ of candidate code-delay values is chosen and used to measure corresponding "correlation-lag" values $\{I_{\hat{\tau}_i} : i = 0, \dots, N - 1\}$, which are then inserted into a maximum-likelihood formula for determining $\hat{\tau}$.

The ranging codes are binary ± 1 -valued signals having a symbol-interval time we denote by t_0 . As explained in a companion article [3], the transmitted waveform is either a Boolean function of several pseudonoise [4] binary sequences running in parallel (the so-called composite-code approach [1]) or a sequence of binary square waves of increasingly short periods (the sequential-code approach). For historical reasons, these were dubbed " τ " and " μ " methods, respectively.

The first planetary ranging system [5] utilized a combined-component-code uplink scheme that required 77 receiver correlations, but only had 2 channels. It was used on Mariner missions from 1969 through 1973, and was then replaced by sequential-component-code machines [1] because of their superior ranging acquisition time advantage of some 12 dB with only 2 correlation channels, at some extra complexity in receiver code switching and housekeeping logic, and at a modest 1.75–2.75 dB loss factor from the optimum matched filter performance.

The acquisition-time advantage of the μ system came from the feasibility of building ranging correlator channels for each of the needed τ_i of the μ code (*viz.*, 2). The τ scheme would have required 77 such channels, a need without cost justification in that era. The merits of the various transmitter codes and receiver detection schemes are adequately treated elsewhere [1], [4], [6], and are not further discussed here.

For determining the phase of incoming symbol transitions, a "clock" component, or period-2 code, is transmitted in both τ and μ systems. The receiver correlation delays are chosen to be $\hat{\tau}_0 = 0$ and $\hat{\tau}_1 = t_0/2$, with $N = 2$.

For a pseudonoise-sequence code [6] of period p bits, the candidate delay values are $\hat{\tau}_i = it_0/k$ for $i = 0, \dots, N - 1$, with $N = kp$. The integer k may be 1 or 2 or more, depending on the transmitter encoding and the method used to determine the received symbol-transition phase, i.e., the code clock delay.

In any case, in order to compute the N values required, either N correlator channels are required to calculate the $I_{\hat{\tau}_i}$ in parallel, or some lesser number may be used serially, but

thereby increasing the acquisition time (equivalently, lowering the effective signal-to-noise ratio).

II. Conventional Ranging Detectors

Conventional DSN ranging correlators are made up of analog and digital hardware and software that mirror the direct calculation of the integral above. Separation into analog and digital portions derives from the following transformation of the correlation integral:

$$I_{\hat{\tau}_i} = \sum_{j=0}^{Mpk-1} c_{j-i} \int_0^{t_0/k} y\left(\frac{t+jt_0}{k}\right) dt \quad (3)$$

where the integration time T is assumed to be a multiple of the period, the code delay $\hat{\tau}_i$ is assumed to be a multiple of the fractional symbol interval, and the coefficients c_j represent the ± 1 code symbol values,

$$T = Mpt_0 = Mpkt_0/k$$

$$\hat{\tau}_i = it_0/k$$

$$c_j = c_m \text{ when } \lfloor j/k \rfloor \equiv \lfloor m/k \rfloor \pmod{p}$$

$$c_{kj} = c(jt_0) \text{ for } j = 0, \dots, p-1$$

The integral thus represents the sum of integrations of the incoming signal over fractional code symbol intervals, multiplied by the appropriate code symbol values over those intervals. One such sum is required for each channel delay $\hat{\tau}_i$, and all multiply's and add's occur in parallel with each other.

The conventional design of a multi-channel range detector is depicted in Fig. 2. Both the old τ and later μ machines used this design, with 2 physical correlator channels. Integration over fractional symbol times is performed by analog hardware, and only one such integrator is needed. The integration value is sampled and converted to a digital number, multiplied by the N delayed code ± 1 stream values, and added to the contents of N accumulators, once each fractional symbol time.

Since k consecutive c_{j-i} values are equal, these sample values may be preaccumulated before multiplications, if desired. This, in fact, is done in the current DSN Spacecraft Ranging System. However, the high-speed logic required for processing the pre-accumulation and subsequent parallel multiply-and-add operations contributes significantly to the ranging assembly cost.

The advent of fast, custom-made, Very Large Scale Integration (VLSI) components has made it possible, albeit still moderately costly, to build many more correlator channels using the same form of digital design as shown in Fig. 2, so that all the needed correlation-delay values can be accumulated digitally in parallel [3].

As a case in point, the ranging uplink design for the Phobos Probe mission uses a pseudonoise code of period 2047, transmitted at about 1.2×10^6 symbols per second. A ranging correlator design using the conventional approach, using only $k = 2$ samples per code symbol, would require nearly 5×10^9 multiply-and-add operations per second. In fact, a preliminary design using the conventional approach was made; it required a massively parallel pipelined architecture utilizing the fastest available memory circuits, and still had to resort to 1-bit quantization of the input signal in order to simplify the multiplier design, with a corresponding 2 dB loss in signal-to-noise performance.

The economics and technology requirements of this conventional approach thus place limits on the code period and performance that can reasonably be expected.

III. The Simplified Digital Correlator

In this section, we note that the correlation integral may be further transformed to reverse the order of digital accumulation and multiplication in calculating the various needed measurements. The advantage of this reversal, as we shall see, is that it reduces the amount of high-speed digital logic and custom-VLSI chips needed, removes almost all of the ranging code dependency from the ranging demodulator hardware, and makes possible the design of a general-purpose ranging demodulator capable of accumulating thousands of correlation-lag values.

By making use of the periodicity of the receiver code, the correlation integral may be further transformed into the equation

$$I_{\hat{\tau}_i} = \sum_{j=0}^{pk-1} c_{j-i} \sum_{m=0}^{M-1} \int_0^{t_0/k} y \left(t + \frac{(j + mpk)t_0}{k} \right) dt \quad (4)$$

or merely

$$I_{\hat{\tau}_i} = \sum_{j=0}^{pk-1} c_{j-i} A_j \quad (5)$$

where A_j represents the accumulation of integrate-and-dump values,

$$A_j = \sum_{m=0}^{M-1} \int_0^{t_0/k} y \left(t + \frac{(j + mpk)t_0}{k} \right) dt \quad (6)$$

The revised correlator depicted in Fig. 3 utilizes pk accumulators for the $\{A_j, j = 0, \dots, pk - 1\}$ values. Note that each integrator sample output is added only into one of the accumulator channels, switched by the accumulator index j , each t_0/k seconds. Because of this simplification, only one adder is required for all the accumulators corresponding to this receiver code, as shown in Fig. 4. This functionally saves $pk - 1$ multiply-and-add logic circuits (4093 of them for the Phobos Probe mission).

For each integrator sample, the corresponding accumulator is fetched, added to the sample, and restored into memory. Each accumulator is only accessed once each pt_0 time interval. Using this technique, the redesigned Phobos Probe ranging correlator requires only 2.4×10^6 additions per second. As a result, conventional Random Access Memory circuitry can be used to hold the values.

Moreover, the accumulators are in jeopardy of overflow only $1/p$ as much as in the conventional design. Consequently, each accumulator can be shorter by $\log_2(p)$ bits than those of the conventional design. This represents an additional logic savings of $pk \log_2(p)$ bits, or about 45,000 bits (5600 bytes) in the Phobos Probe ranging correlator storage.

The address generator is merely a counter clocked at the code fractional symbol rate and reset at the beginning of each period of the receiver code. This is the only code-dependent signal entering the digital portion of the design. Except for the number of lag-value accumulators, the digital portion of the correlator assembly is completely independent of the receiver code components. No receiver coder hardware is necessary. Conventional deep-space ranging correlators required hardware code generators for each of the different receiver code components.

Code multiplications are not made until after the complete accumulation of A_j values has been read into the computer of the ranging assembly. Then the same A_j set serves to calculate all of the $I_{\hat{\tau}_i}$ for $i = 0, \dots, N - 1$. The vector \mathbf{I} containing the $I_{\hat{\tau}_i}$ is related to the vector \mathbf{A} of accumulator values A_j by the equation

$$\mathbf{I} = \mathbf{CA} \quad (7)$$

where \mathbf{C} is the $N \times kp$ matrix of binary code values, $c_{i,j} = c_{j-i}$. Since the receiver codes are stored solely in the computer memory as \mathbf{C} , since the vector \mathbf{I} is computed separately from the accumulation process, and since this computation only needs to take place infrequently, there is a greater degree of flexibility and generality in the simpler design than existed in the previous ranging assemblies.

IV. Conclusion

This article has presented a design simplification for the digital hardware design of a deep-space binary-code ranging system. The simplification is significant in that it

- (1) Makes use of easily obtainable RAM storage, one memory location for each correlation lag to be accumulated, accessed serially, rather than special VLSI devices or high-speed-logic circuits accessed in parallel;
- (2) Reduces the high-speed digital logic requirements to a single sample accumulator and adder, regardless of the

number of correlation lags computed, representing a savings of $pk - 1$ multiply-and-add circuits;

- (3) Reduces the number of bits required by each accumulator by $\log_2(p)$ bits each, for a savings of $pk \log_2(p)$ bits for the entire correlator;
- (4) Makes the analog and digital portions of the system independent of both the transmitter and receiver codes, except for timing signals;
- (5) Removes the need for receiver coders (simple period counters will do); and
- (6) Stores the receiver code component(s) as a vector in the ranging computer, where the accumulated correlation-lag values are computed by a single simple matrix multiplication, infrequently calculated.

Using this approach, the new Phobos Probe mission ranging system design uses no special components, is smaller, is easier to design and maintain, and does not need to 1-bit quantize the input signal. The details and particulars of the correlator design will be the subject of a subsequent article.

References

- [1] J. H. Yuen (ed.), *Deep Space Telecommunications Systems Engineering*, JPL Publication 82-76, Jet Propulsion Laboratory, Pasadena, California, pp. 123-178, 1972.
- [2] W. B. Davenport and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*, New York: McGraw-Hill, 1958.
- [3] R. C. Tausworthe, "Tau Ranging Revisited," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 318-323, November 15, 1987.
- [4] R. C. Tausworthe, "Optimum Ranging Codes," *Trans. of the IEEE PTG-SET*, vol. 10, no. 1, pp. 19-30, March 1964.
- [5] R. C. Tausworthe, "Ranging the 1967 Mariner to Venus," in *Proc. of IEEE National Convention*, New York, pp. 294-295, March 1967.
- [6] S. W. Golomb, et al., *Digital Communications with Space Applications*, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964.

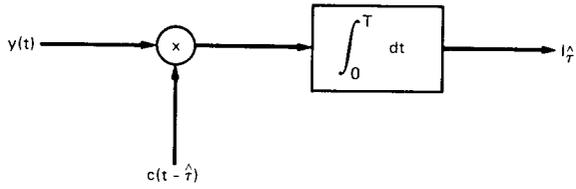


Fig. 1. Ranging correlator channel functional design

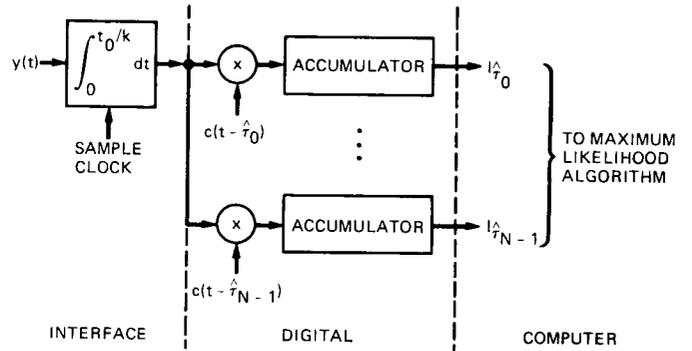


Fig. 2. Conventional multi-channel ranging correlator design

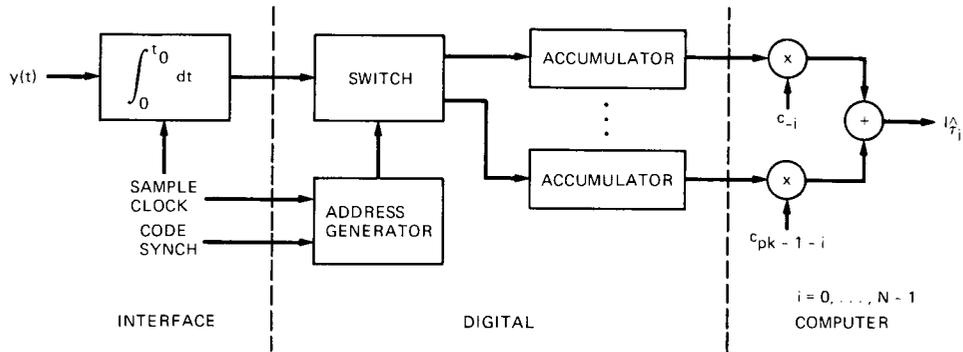


Fig. 3. Revised multi-channel ranging correlator design

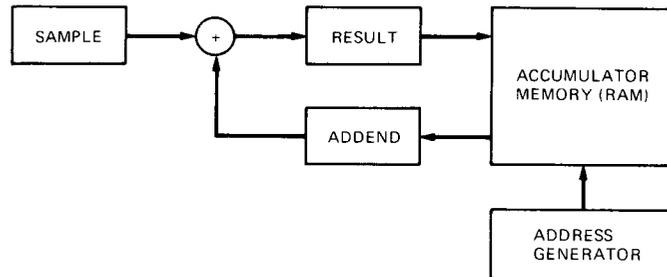


Fig. 4. Ranging correlator digital processor

S/3-17

128229

TDA Progress Report 42-92

TP

N88-18787

October-December 1987

Frame Synchronization Performance and Analysis

C. S. R. Aguilera, L. Swanson, and G. H. Pitt III
Communications Systems Research Section

This article describes the analysis used to generate the theoretical models showing the performance of the frame synchronizer for various frame lengths and marker lengths at various signal-to-noise ratios and bit error tolerances.

I. Introduction

The telemetry data stream from a deep space mission, usually many hundreds of thousands of bits long, is divided into units called "frames." Each frame is introduced by a "marker" (generally 32 bits long, but not always) announcing the frame boundary. Until now, the Deep Space Network (DSN) has sent the data stream to the projects without regard to frame boundaries.

The DSN is beginning to install "frame synchronizers" which look for markers and divide the data into frames. Beginning with Magellan, the DSN will deliver data to the projects already divided into frames. The currently planned configuration is for these frame synchronizers to work on decoded information bits; therefore, they must follow the Viterbi decoder (see Fig. 1). See [1] for a discussion of the relative merits of frame synchronization before and after Viterbi decoding.

The technical problems of dividing a noisy data stream into frames are probabilistic, depending on the channel. When the data stream arrives, a search for the marker is conducted. However, because of channel noise, the marker might be corrupted;

on the other hand, the actual data within the frame might include 32 bits that are exactly the same as the marker. Thus, when looking at the data, the frame synchronizer needs to decide which bits most likely represent the marker in order to decide where the frame begins.

II. The Frame Synchronizer

Before describing the Frame Synchronizer Subassembly (FSS), what is meant by "finding" the marker must be established.

Since it is likely that a marker may contain errors due to normal deep space communications noise, a match is "found" when a string of bits disagrees with the marker in no more than T places, where T is the Bit Error Tolerance (BET) threshold [2]. If more than T errors are made in the frame sync word, then it will be "missed." The "best match" is the string of decoded bits disagreeing with the marker in the fewest places. During the search for the frame sync marker (see below), if the minimum error detector (MED) is enabled, then the best match identified in one frame length is used as the marker. Otherwise, the first match is used. (If errors in the

decoded bit error stream were independent, then the best match would most likely be the marker. See [2] for another method taking Viterbi decoder error statistics into account.)

After a possible marker is identified, the data stream is examined at a point one frame length away for another marker. If a match is found, then the marker is considered verified once. This process may be repeated for various numbers of verifications.

The frame synchronizers in the DSN use the following algorithm (see Figs. 2 and 3), given $T = \text{BET}$, $K =$ the number of verifies, and $N =$ the number of flywheels:

- (1) **SEARCH:** Search the data stream until a possible marker is found with fewer than T disagreements with the marker. If the MED is enabled, continue the search for one complete frame length.
- (2) **VERIFY:** Examine the data stream at a point one frame length away for another marker. If the match is within T of the marker, continue the verification; otherwise return to step 1. If K verifications are successful, then "frame in lock," or sync, is established.
- (3) **LOCK:** Continue testing the marker. If more than T errors are found, enter the flywheel mode. FSS has sync while in this mode.
- (4) **FLYWHEEL:** Determine whether the sync is lost by testing up to N consecutive frames. If the marker is found within the N frames, then sync is reestablished and the FSS returns to step 3. Otherwise, return to step 1.

III. Analysis and Probabilities

In analyzing the probability of finding the sync marker, only two cases were considered: (1) the probability of frame acquisition within $F = 4$ frames with K verifications ($K = 3, 2, \text{ or } 1$) and the MED enabled; and (2) the probability of frame acquisition within F frames ($F = 1, 2, 3, \text{ or } 4$) with no verifications and the MED enabled.

Using data from Table C-1 of [3] (Viterbi Decoder Burst Statistics: 3233013 [7, 1/2] Convolutional Code), error distribution data was generated by simulation and used to determine the probabilities for the number of errors made by the Viterbi decoder during a 32-bit span of bits, and thus the probabilities of different numbers of disagreements between the decoded marker and the true marker. Below, D is used for the number of disagreements between the decoded marker and the true marker, so $P(D = j)$ is the probability that the Viterbi decoder makes j errors in the marker.

The binomial probability distribution with $p = 1/2$ was used to determine the probabilities for the number of disagreements between 32 random bits and the marker. R is the number of disagreements between 32 random bits and the markers, so

$$P(R = j) = \binom{32}{j} 2^{-32}$$

The analysis below assumes that in all cases of "ties," i.e., the same number of disagreements in two places, the frame synchronizer makes the wrong decision. This means that actual performance will be slightly better than that predicted below.

Variable	Typical values
$T =$ threshold	$0 \leq T \leq 32$
$L =$ frame transport length	$L = 5120, 6720, 480, 16320$
$F =$ number of frames	$F = 1, 2, 3, 4$
$K =$ number of times sync is verified after finding the initial marker	$K = 1, 2, 3$

A. Probability of Frame Acquisition Within $F = 4$ Frames With K Verifications

With $K = 3$ verifications and $F = 4$ frames, the marker must be found correctly in each of the frames:

$$P(\text{sync found correctly and verified 3 additional times}) \\ = P(\text{sync found correctly in first frame and verified 3 times})$$

Assuming independence between frames, this is

$$P(\text{sync found correctly}) \cdot [P(\text{verified})]^3 \\ = \sum_{j=0}^T P(D = j) \cdot [P(R > j)]^L \cdot [P(D < T)]^3$$

A large assumption is implicit in the use of $[P(R > j)]^L$ for the probability that the marker is never found in random data. Finding the marker in different places in random data is not independent unless the places are at least a marker length apart. The actual probability that the marker is found in random data is at least $1 - 32(1 - [P(R > j)]^{L/32})$, which is close to $[P(R > j)]^L$ for small j . As j increases, these quantities begin to differ, but the contribution to the sum becomes small. An actual calculation of the probability of finding the marker with up to j disagreements somewhere in a frame is

analytically intractable, and approximation by simulation uses a great deal of computer time (see [1]).

The probability that fewer than T errors are found in the sync word 3 times in a row is $[P(D < T)]^3$, assuming independence between different frames.

With $K = 2$ verifications and $F = 4$ frames, sync can be declared correctly in two ways:

$P(\text{sync found correctly and verified 2 additional times})$

$= P(\text{found correctly in 1st frame and verified twice})$
 $+ P(\text{found correctly in 2nd frame and verified twice, not found in 1st frame})$

$= P(\text{found correctly in 1st frame and verified twice})$
 $+ P(\text{found correctly in 2nd frame and verified twice})$
 $\cdot P(\text{not found anywhere in 1st frame})$

$= P(\text{found correctly and verified twice})$
 $\cdot \{1 + P(\text{not found anywhere in 1st frame})\}$

$$= \sum_{j=0}^T P(D=j) \cdot [P(R>j)]^L \cdot [P(D<T)]^2$$

$$\cdot \left\{ 1 + P(D>T) \cdot [P(R>T)]^L \right\}$$

With $K = 1$ verification and $F = 4$ frames, sync can be declared correctly in five ways:

$P(\text{sync found correctly and verified 1 additional time})$

$= P(\text{found correctly in 1st frame and verified once})$
 $+ P(\text{found correctly in 2nd frame and verified once, not found in 1st frame})$
 $+ P(\text{found correctly in 3rd frame and verified once, not found in first two frames})$
 $+ P(\text{found correctly in 3rd frame and verified once, found correctly in 1st frame, but not verified})$
 $+ P(\text{found correctly in 3rd frame and verified once, found incorrectly in 1st frame, but not verified})$

$= P(\text{found correctly and verified once})$
 $\cdot \{1 + P(\text{not found anywhere in 1st frame})$
 $+ P(\text{not found anywhere in 1st and 2nd frames})$
 $+ P(\text{found correctly in 1st frame, but not verified})$
 $+ P(\text{found incorrectly in 1st frame, but not verified})\}$

$$= \sum_{j=0}^T P(D=j) [P(R>j)]^L \cdot [P(D<T)]$$

$$\cdot \left\{ 1 + P(D>T) \cdot [P(R>T)]^L \right.$$

$$+ [P(D>T) \cdot [P(R>T)]^L]^2$$

$$+ \sum_{j=0}^T P(D=j) \cdot [P(R>j)]^L \cdot P(D>T)$$

$$\left. + P(D>T) \cdot (1 - [P(R>T)]^L) \cdot P(R>T) \right\}$$

Two sets of graphs can be generated from the data. One set is a function of SNR, the other a function of BET.

B. Probability of Frame Acquisition Within F Frames With No Verifications

With $F = 1$ frame:

$P(\text{sync found correctly in 1 frame})$

$$= \sum_{j=0}^T P(D=j) \cdot [P(R>j)]^L$$

With $F = 2$ frames:

$P(\text{sync found correctly in 2 frames})$

$= P(\text{found correctly in 1st frame})$
 $+ P(\text{found correctly in 2nd frame, not found in 1st frame})$

$$= \sum_{j=0}^T P(D=j) \cdot [P(R>j)]^L$$

$$\cdot \left\{ 1 + P(D>T) [P(R>T)]^L \right\}$$

With $F = 3$ frames:

$P(\text{sync found correctly in 3 frames})$

$= P(\text{found correctly in 1st frame})$
 $+ P(\text{found correctly in 2nd frame, not found in 1st frame})$
 $+ P(\text{found correctly in 3rd frame, not found in first two frames})$

$$\begin{aligned}
&= \sum_{j=0}^T P(D=j) \cdot [P(R>j)]^L \\
&\quad \left\{ 1 + P(D>T)[P(R>T)]^L \right. \\
&\quad \left. + [P(D>T)[P(R>T)]^L]^2 \right\}
\end{aligned}$$

With $F = 4$ frames:

$$\begin{aligned}
&P(\text{sync found correctly in 4 frames}) \\
&= P(\text{found correctly in 1st frame}) \\
&\quad + P(\text{found correctly in 2nd frame,} \\
&\quad \text{not found in 1st frame}) \\
&\quad + P(\text{found correctly in 3rd frame,} \\
&\quad \text{not found in first two frames}) \\
&\quad + P(\text{found correctly in 4th frame,} \\
&\quad \text{not found in first three frames})
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j=0}^T P(D=j) \cdot [P(R>j)]^L \\
&\quad \cdot \left\{ 1 + P(D>T)[P(R>T)]^L \right. \\
&\quad \left. + [P(D>T)[P(R>T)]^L]^2 \right. \\
&\quad \left. + [P(D>T)[P(R>T)]^L]^3 \right\}
\end{aligned}$$

IV. Frame out of Lock

The probability of Frame out of Lock occurring at least N consecutive times is as follows:

Frame Error Distribution: $P(D > T)$

Frame out of Lock: $[P(D > T)]^N \quad N = 1, 2, 3, 4, 5$

V. Numerical Results

Figures 4 through 6 generated by the models represent the two main operating FSS modes: (1) "Frame Acquisition," comprising the Search and Verify modes; and (2) "Frame in Lock," comprising the Lock and Flywheel modes.

Many curves, using different parameters, were generated, but only three representative curves will be included here. Figure 4 shows the probability of frame acquisition within $F = 4$ frames with two verifications versus bit signal-to-noise ratios (SNR or E_b/N_0) for various BETs. Figure 5 shows the probability of frame acquisition within $F = 4$ frames with SNR = 2.1 dB versus the BET for various numbers of verifications. Figure 6 shows the probability that the marker is corrupted by more than BET errors $N = 2$ times in a row versus the bit SNRs for various BETs.

Figures 4 and 5 give probabilities of "Frame Acquisition," while Fig. 6 gives probabilities of "Frame out of Lock," the opposite of "Frame in Lock." It is easy to see that the number of combinations of parameters is far too great to allow a full set of graphs to be included here.

All of these graphs are based on the "geometric model" for Viterbi decoder burst errors, and on the statistics in [3].

VI. Conclusions

Several parameters may be chosen when operating the Frame Synchronizer Subassembly. Some of these are T , the bit error threshold which specifies the maximum number of differences allowable when declaring a set of bits the frame sync marker, K ; the number of verifications needed before declaring the stream of bits in lock; and N , the number of flywheels to go through before declaring a stream out of lock. The operator must also set the length of the marker and the length of the frame.

There is no obviously best choice for T , K , and N because each one gives a trade-off between the probability of incorrectly declaring lock and incorrectly not declaring lock; the merits of the two competing factors must be weighed. However, the graphs at the end of this report should help in making that decision by giving probability estimates in several cases. The results of this study are summarized in the graphs.

References

- [1] L. Swanson, *A Comparison of Frame Synchronization Methods*, JPL Publication 82-100, Jet Propulsion Laboratory, Pasadena, California, 1982.
- [2] M. Shahshahani and L. Swanson, "A New Method for Frame Synchronization," *TDA Progress Report 42-90*, vol. April-June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 111-123, August 15, 1986.
- [3] R. L. Miller, L. J. Deutsch, and S. A. Butman, *On the Error Statistics of Viterbi Decoding and the Performance of Concatenated Codes*, JPL Publication 81-9, Jet Propulsion Laboratory, Pasadena, California, 1981.

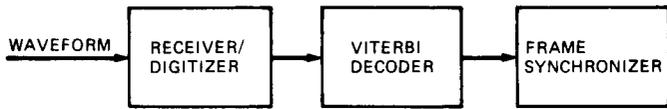


Fig. 1. Proposed telemetry system, including the FSS

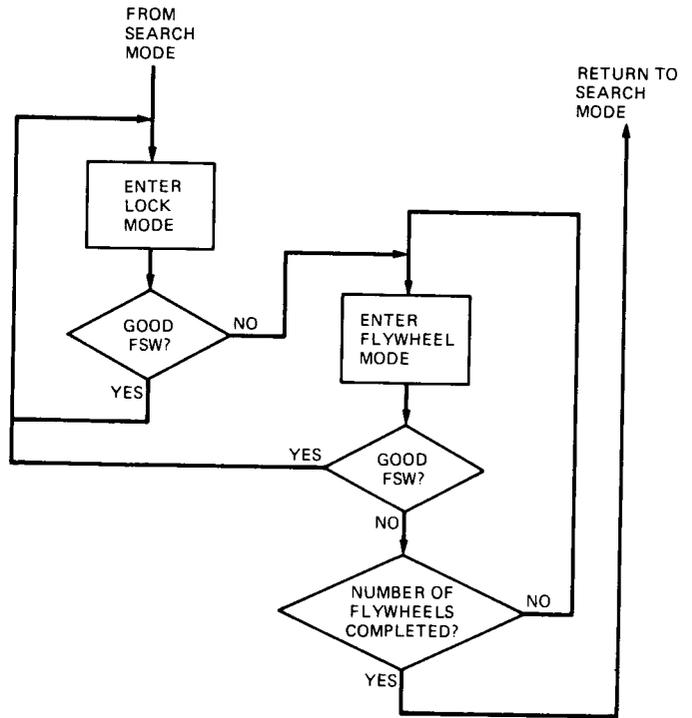


Fig. 3. Flowchart for lock and flywheel modes

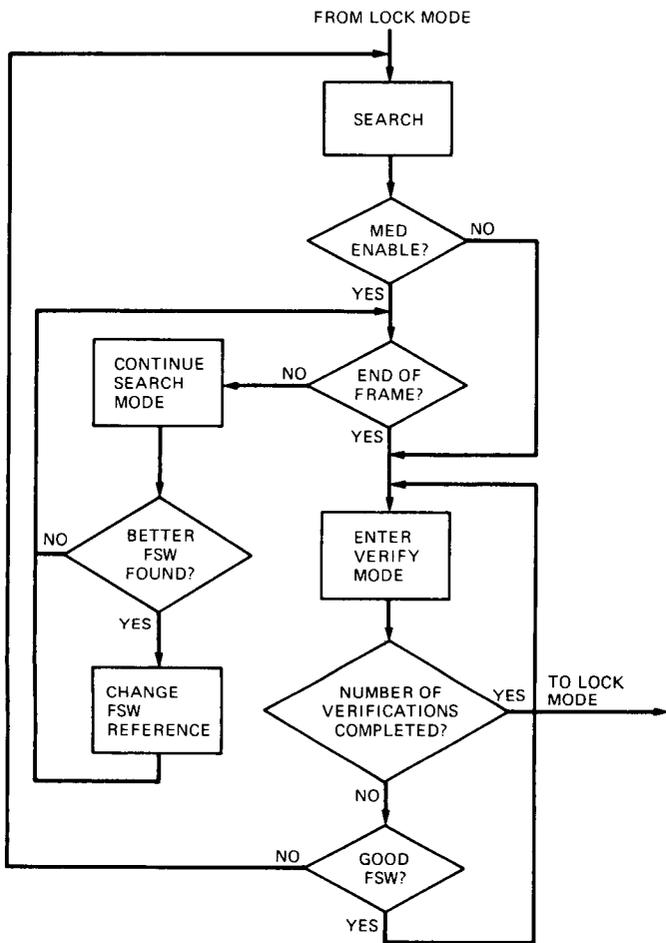


Fig. 2. Flowchart for search and verify modes

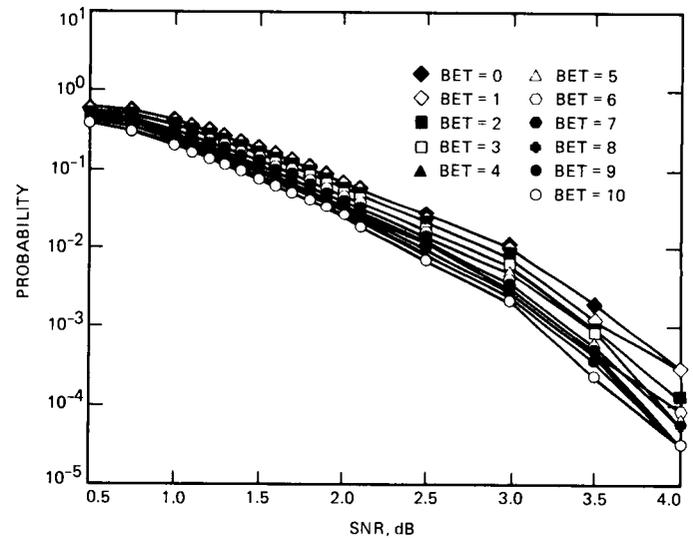


Fig. 4. Frame acquisition within a fixed time with a fixed number of verifications. This chart shows 1 minus the probability that correct sync is found within four frames with two verifications and a frame transport length of 5120 (from computations using the analysis above) versus the SNR.

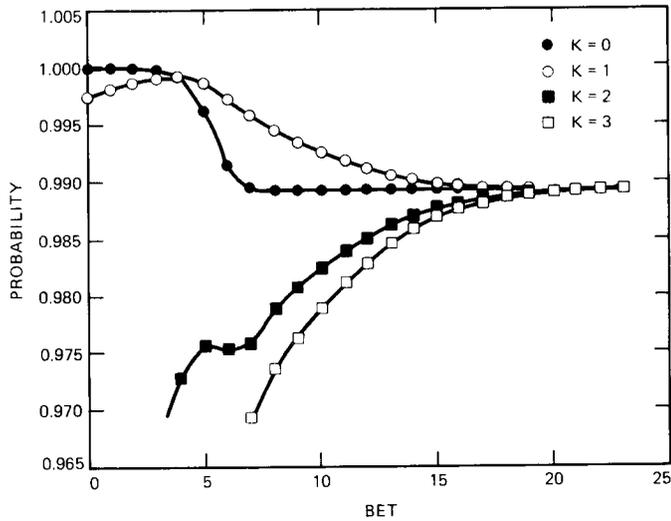


Fig. 5. Frame acquisition within a fixed time with a fixed SNR. This chart shows the probability that correct sync is found within four frames at 2.1 dB and a frame transport length of 5120 (from computations using the analysis above) versus the bit error threshold. K is the number of verifications.

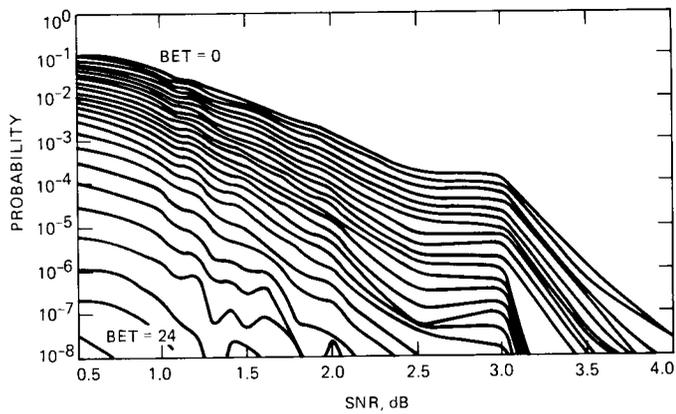


Fig. 6. Frame out of lock with a fixed number of flywheels. This chart shows 1 minus the probability that correct sync is lost at a given time with two flywheels ($1 - P[D > \text{BET two times in a row}]$) versus the SNR.

A Noise-Adding Radiometer for the Parkes Antenna

T. J. Brunzie

Radio Frequency and Microwave Subsystems Section

A new design for a noise-adding radiometer will be included as part of the reimplementation of the Parkes antenna microwave front end. Designed as an aid for antenna calibration, the Parkes NAR will support the Voyager-Neptune encounter in 1989.

I. Introduction

During the upcoming Voyager-Neptune encounter, the 64-meter Australian National Radio Astronomy Observatory at Parkes will once again assume the role of a DSN tracking station as part of the Parkes-Canberra Telemetry Array (PCTA). As with the earlier Uranus encounter, this requires outfitting the antenna with DSN-compatible hardware, ranging from microwave feedhorn to telemetry receiver. Although much of the hardware used during the previous encounter will be reimplemented for the Neptune encounter, the European Space Agency (ESA)-designed front-end monitor and control system will be replaced with a new system designed by JPL.

Among the ESA features that will be duplicated by the new Parkes Front-End Controller (FEC) is a Noise-Adding Radiometer (NAR), a device used to measure antenna system temperature. Its operation is based on the fact that receiver noise power is directly proportional to temperature. Thus, measuring the relative increase in noise power due to the presence of a calibrated thermal noise source allows direct calculation of system temperature. Although it is intended primarily to aid pre-pass antenna pointing calibration procedures, the Parkes NAR will also be capable of monitoring system temperature during telemetry tracks without significant degradation through the use of low-noise diodes.

II. System Configuration

The NAR implemented as part of the monitor and control for the Parkes antenna front end will consist of two basic subsystems: an array of noise diodes, located in the aerial cabin, for injecting noise into the system, and a precision power meter, forming part of the FEC computer, for measuring noise power. The noise diode assemblies are DSN standard equipment—duplicates of the diode ovens and power supplies that form part of the DSN Precision Power Monitor (PPM) assembly. The Digital Power Meter (DPM), a new design, is a functional replacement for the PPM's square-law diodes, employing digital signal processing techniques for noise measurement.

The heart of the system is the Parkes FEC, an 86/14-based multibus computer containing the DPM and configured for monitor and control of the front-end microwave electronics (Fig. 1). The FEC's tasks will include remote control of the noise diode assemblies and operation of the DPM for performing NAR temperature measurements.

Two noise diode assemblies will be provided for the Parkes antenna, one for each of the two X-band receive chains. Each assembly consists of a noise diode oven and associated power supply. Each oven contains three diodes, providing noise tem-

peratures of approximately 0.25, 0.5, 1, 2, 4, 8, and 50 K (defined at the maser input). The diodes are controlled through their power supply assemblies, with relays being used to select the amount of diode current (allowing three noise levels per diode); a fourth TTL signal is used to modulate the diode on and off.

Each power supply assembly is monitored and controlled through 21 digital I/O lines, consisting of relay closures, closure sense, and diode modulation input. Both assemblies will be operated through an HP 3488A switch/control unit containing three HP 44474A digital I/O cards (16 channels per card). A coaxial cable run directly from the FEC will supply the modulation control signals.

Noise power measurements will be made at the inputs of the Parkes telemetry receiver. Each of the two 320-MHz RF signals will be split 3 dB in the receiver signal select drawer and then fed directly to the DPM in the FEC. The DPM consists of three multibus PC boards under FEC control that sample the inputs and accumulate measured power values. An averaged output is read over the bus by the FEC 86/14 CPU. Diode control, noise measurement, system temperature computation, and an analog output will all be handled by the CPU, with results included in FEC status displays.

III. Parkes NAR Operation

Noise-adding radiometers operate by periodically injecting a known quantity of noise into an antenna front end and then measuring the resulting increase in system noise power at the receiver. System temperature measured this way can be used for antenna pointing calibration (star tracking), system performance history, and spacecraft power measurement. Because these tasks usually require a continual stream of system noise temperature data, the Parkes NAR will measure noise levels and compute temperature repeatedly, and will return its results to the antenna pedestal in both digital and analog form.

The Parkes NAR operates by selecting a noise diode and diode current needed to achieve a desired additive noise temperature and then modulating the diode on and off while measuring power at the Parkes telemetry receiver. Two noise power measurements are needed to calculate system temperature: one while the diode is on and one while it is off. A simple calculation based on the diode temperature and the two measured values yields the unknown system temperature.

Although noise power measurements for NAR operation are ideally taken from the telemetry receiver IF, it will not be practical to do so in the case of the Parkes NAR. Measurement using the PCTA receiver 70-MHz IF would require operating

the receiver subsystem in addition to the FEC. However, by taking measurements at the receiver 320-MHz inputs, the NAR becomes independent of PCTA operation. As seen in Fig. 1, the use of 3-dB power splitters within the receiver signal select drawer provides the needed signals.

The measurement process begins with control of the noise diode assemblies. This will be done using an HP 3488A switch control unit, rather than a PPM Noise Diode Controller assembly. Operated remotely over the IEEE-488 GP-IB bus, the HP 3488A sets the power supply relays that control diode current and monitors the relay closures to verify proper settings. Diode modulation is controlled directly by the DPM in coordination with noise power measurements.

Software operation consists of selecting the desired noise diode temperature (approximately 0.25, 0.5, 1, 2, 4, 8, or 50 K), selecting the desired measurement rate, and programming the type of analog output desired. System temperature readings will then be available through either status polls or time-dependent graphs. The entire process is initiated and timed by the FEC CPU, which includes in its loop a routine to drive a digital-to-analog converter with the results of the calculations. This analog output will be fed back to the antenna calibration facility in the antenna pedestal.

Within the FEC, the DPM (Fig. 2) performs noise power measurements in a fixed bandwidth of 320–340 MHz by averaging the square of a large number of sampled noise voltages. Under CPU control, an RF switch and a programmable attenuator select the input channel and adjust the noise level to a fixed gain. The attenuator not only has sufficient range for expected noise level variations but can also adjust for inputs from the antenna's ambient load, and will be used for making Y-factor measurements. Next, an on-board local oscillator fixed at 330 MHz mixes the input down to base-band, which is then low-pass filtered at 10 MHz. An 8-bit analog-to-digital converter generates the digital samples of the noise voltages, which are then fed at 20 MHz to a 34-bit-wide multiplier/accumulator for squaring and averaging. The resulting total noise power value is read from the board directly by the CPU.

Timing and control for each measurement is handled by the DPM; the CPU is needed only to read the resulting averaged noise power and to reinitiate the measurement process. The measurement time is variable and can be controlled by the CPU; measurement rate is determined by how often the CPU initiates the process. An interrupt and a status flag are available to signal the CPU each time the process is completed.

In order to achieve both a short measurement time and high accuracy, a total of 2^{18} samples (nominally) are taken

at a rate of 20 million samples/second, yielding a 13-ms measurement time and a sampling accuracy of 0.2 percent. (Nyquist sampling theory does not apply to this case, since only noise power is of interest, not the ability to reconstruct waveforms.) Total time overhead includes an additional 2 ms "dead time" between the switching of the diodes and the start of each measurement; this gives the system a chance to settle and allows the CPU ample time to compute the results and restart the process. (Figure 3 diagrams the software loop timing coordinating FEC operation and NAR measurements.) Given that two noise measurements must be made to compute system noise temperature, with four measurements between switching, the nominal sampling rate for these T_{op} measurements is then twice $(4 \times 13 + 2)$ ms, or 9.5 Hz. Resolution is controlled through averaging of the T_{op} samples.

Periodically during NAR operation, the FEC inserts an extra measurement in the loop for determining DC offset. This is done by computing an average of noise voltage samples rather than the square of samples. This offset is used to eliminate the DC component from the total noise power measurement, yielding a purely AC noise power figure for computing T_{op} .

One advantage to implementing the NAR within the Parkes FEC is that the system can be either operated in a stand-alone mode in conjunction with the other front-end equipment or automated within the entire PCTA Receiver/Combiner subsystem. The range of low-noise diodes provided in the PPM diode assembly allows the use of the NAR during telemetry tracking with minimal degradation of telemetry data. Integrated system operation will be available to the Parkes receiver for real-time temperature measurement, and to CDSCC for remote operation and/or monitoring of system performance.

A second advantage provided by the FEC/DPM is the ability to coordinate with the Parkes front-end equipment during test/calibration procedures. The DPM has sufficient range to measure noise power from the antenna's ambient load as well as the cold sky. This makes it possible to reference the ambient load for calibrating the diodes and compensating for system nonlinearities using any one of several techniques [1]. Since the FEC controls the front-end equipment in addition to the NAR, waveguide switching, maser selection, and DPM operation can all be controlled by one program, either with a backup CRT terminal in a stand-alone mode post-pass or automatically as part of a PCTA precalibration configuration control file.

IV. NAR Temperature Calculations

Given the fact that noise power and noise temperature in an antenna system are directly proportional to one another,

two equations can be formed from the presence and absence of a known additive noise source:

$$P_{off} = k(T_{op})$$

$$P_{on} = k(T_{op} + T_d)$$

where

- P_{off} = system noise power with noise diode off, W
- P_{on} = system noise power with noise diode on, W
- T_{op} = operating noise temperature, K
- T_d = noise diode temperature, K
- k = proportionality constant, W/K

Combining these two equations to eliminate k ,

$$\frac{P_{on}}{P_{off}} = \frac{T_{op} + T_d}{T_{op}} = 1 + \frac{T_d}{T_{op}}$$

Rearranging yields

$$T_{op} = \frac{T_d}{\left(\frac{P_{on}}{P_{off}}\right) - 1}$$

Thus, making two noise power measurements using a diode of known temperature allows a direct calculation of system noise temperature insensitive to low-frequency gain changes. It can be seen in the calculation of T_{op} that a sizable difference in noise power would help reduce sensitivity to errors in power measurement. While this can easily be accomplished through the use of large (50 K) noise diodes during antenna calibration, operating the NAR during a Voyager array pass would require the use of small diodes in order to prevent significant telemetry degradation.

An additional problem in present DSN NARs involves nonlinearities in measurements performed by the PPM. The problem lies with the PPM square-law diode detectors not being square-law. The Parkes NAR will attempt to improve noise measurement accuracy by replacing the nonlinear square-law diodes with the new linear DPM.

A third consideration in calculating antenna system noise temperature is resolution. The degree to which any NAR can resolve noise temperature is expressed by the following equation [2]:

$$(\Delta T_{\min}) \text{NAR} = \frac{2T_{\text{op}}}{(tB)^{1/2}} \left[1 + \frac{T_{\text{op}}}{T_d} \right]$$

where

B = detector bandwidth, Hz

t = total integration time, s

In the case of the Parkes NAR, a best-case system temperature of 21.5 K (maser 1, antenna at zenith), the 50-K noise diode for antenna pointing, the DPM bandwidth of 10 MHz, and a requirement on ΔT_{\min} of 0.01 K allow solution of a minimum integration time, t_{\min} :

$$t_{\min} = \frac{4T_{\text{op}}^2}{B(\Delta T_{\min})^2} \left[1 + \frac{T_{\text{op}}}{T_d} \right]^2$$

$$t_{\min} = \frac{4(21.5)^2}{10^7(0.01)^2} \left[1 + \frac{21.5}{50} \right]^2 = 2.64 \text{ s}$$

Total measurement time includes switching "dead time" in addition to total integration time. Similar calculations using a

worst-case T_{op} of 26 K (maser 2, antenna at 25 degree elevation) yield a total integration time of 4.11 seconds. Since the DPM computes a T_{op} value in about 100 ms, a large number of samples would need to be averaged together by the FEC CPU to meet the total integration time. (Figure 4 illustrates the relationship between temperature resolution and integration time for a variety of diode temperatures using a typical case of $T_{\text{op}} = 24 \text{ K}$.)

In addition to averaging T_{op} samples to meet resolution requirements, the FEC will also have the ability to automatically adjust integration time to continually compensate for variations in the computed T_{op} , thereby keeping noise temperature measurement resolution within specification at all times.

V. Conclusion

Although not yet out of the proof-of-concept phase the design of the Parkes noise-adding radiometer has generated enough support for inclusion in the Parkes Front-End Reimplementation Task. A thorough RF analysis has been completed, and plans call for testing a breadboard of the DPM in early 1988. Full-scale system testing with the front-end electronics is scheduled for mid-1988, with delivery of the completed system late in the year.

Acknowledgments

The author would like to thank C. T. Stelzried for providing suggestions on NAR operation and H. R. Buchanan for his help in RF analysis.

References

- [1] C. T. Stelzried, "Non-Linearity in Measurement Systems: Evaluation Method and Application to Microwave Radiometers," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 57-66, November 15, 1987.
- [2] P. D. Batelaan *et al.*, *A Noise-Adding Radiometer for Use in the DSN*, JPL Space Programs Summary 37-65, vol. 2, Jet Propulsion Laboratory, Pasadena, California, pp. 66-69, September 30, 1970.

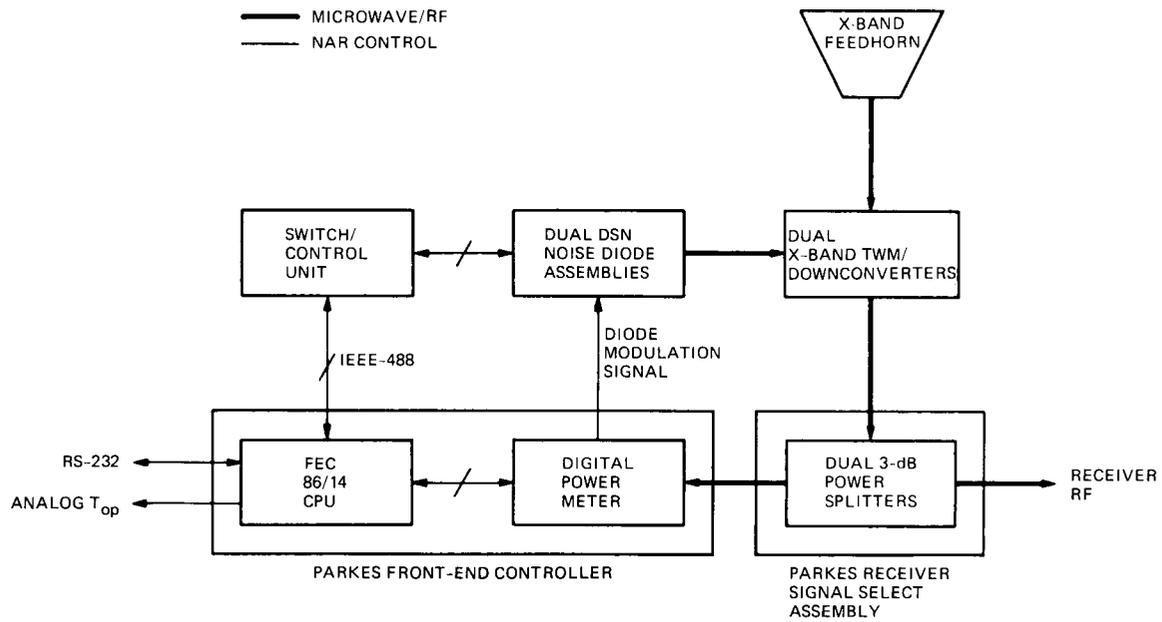


Fig. 1. Parkes noise-adding radiometer block diagram

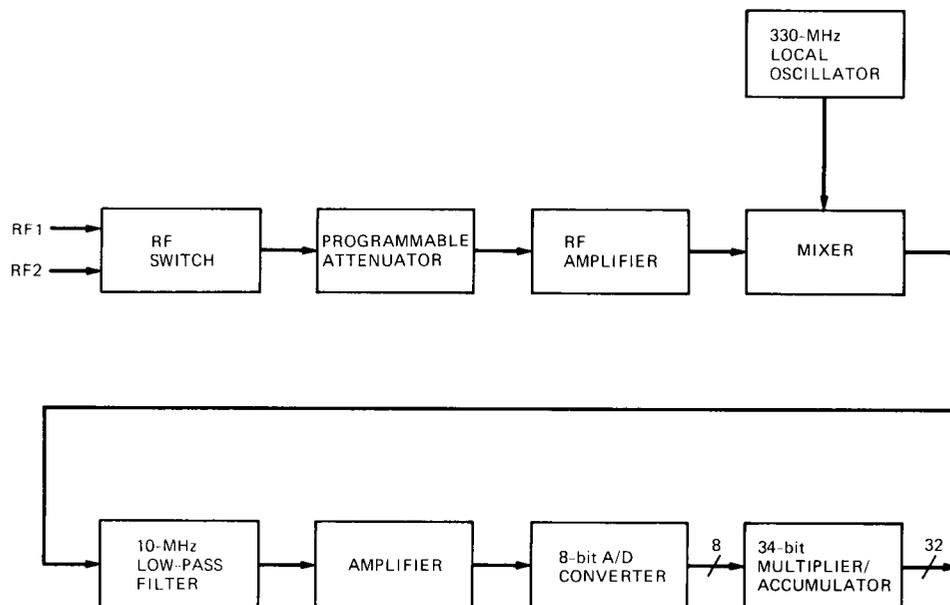


Fig. 2. Digital power meter block diagram

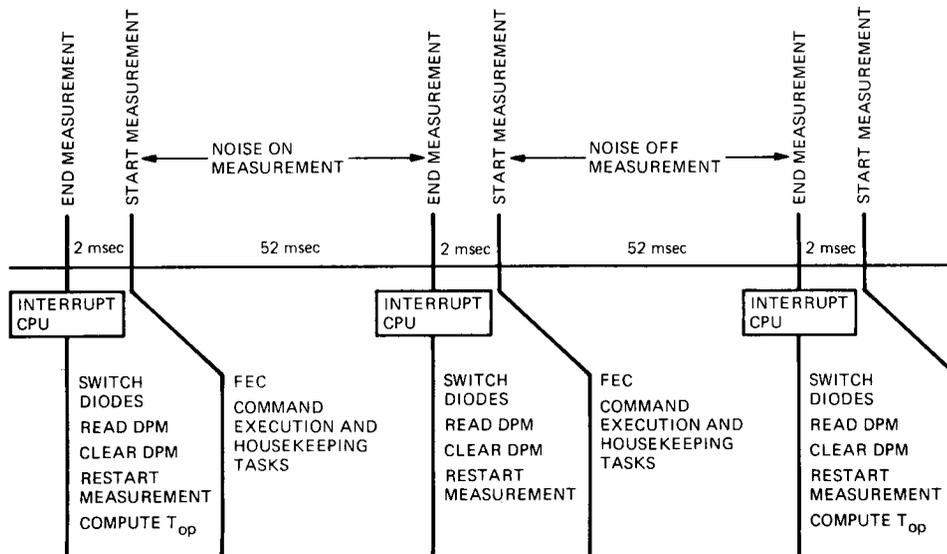


Fig. 3. Front-end controller/NAR loop operation and timing

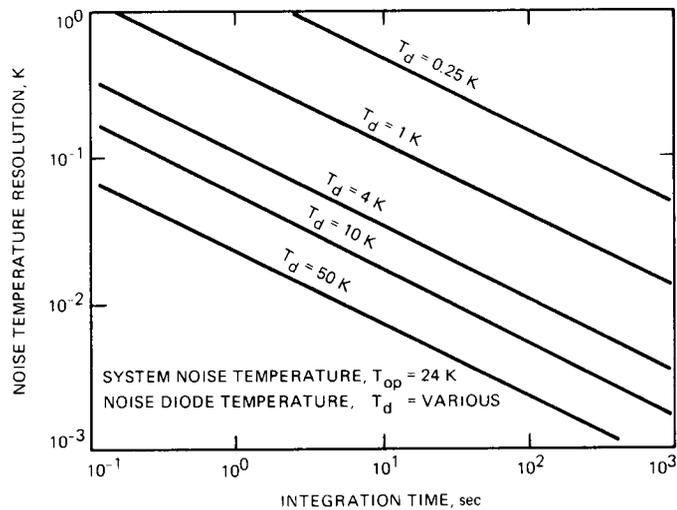


Fig. 4. NAR temperature resolution/integration time relationship

X-Band System Performance of the Very Large Array

J. S. Ulvestad

Tracking Systems and Applications Section

G. M. Resch

TDA Technology Development Office

W. D. Brundage

National Radio Astronomy Observatory
Socorro, New Mexico

The Very Large Array (VLA) is being equipped to receive telemetry from Voyager 2 during the Neptune encounter in 1989. Cryogenically cooled amplifiers are being installed on each of the 27 antennas. These amplifiers are currently a mix of field effect transistors (FETs) and high electron mobility transistors (HEMTs) and exhibit zenith system temperatures that range from 30 K to 52 K. We summarize the system temperatures and aperture efficiencies determined during the past year. The nominal values of the noise diode calibration are compared with derived values made under the assumption of a uniform atmosphere over the array. Gain values are determined from observations of unresolved radio sources whose flux densities are well known. The tests suggest that the completed VLA will have a ratio of gain to system temperature that is approximately 4.4 dB above that of a single 64-m antenna of the Deep Space Network.

I. Introduction

The Very Large Array (VLA) [1] in New Mexico will be arrayed with two or three Goldstone antennas during the Voyager encounter with Neptune and will support 21.6 kilobit/s telemetry from the spacecraft [2]. Initial plans called for equipping the VLA with cryogenically cooled FET amplifiers with an expected system temperature performance of 45–50 K at the Voyager downlink frequency of 8.42 GHz. During the planning phase of this implementation, the Deep Space Network (DSN) Advanced Systems Program and the National Radio Astronomy Observatory (NRAO) jointly began an effort with the General Electric Corporation and with Cornell University to examine the potential of a new type

of transistor—a high electron mobility transistor (HEMT) [3] that might replace the FET and give better performance. This effort has been successful with the result that HEMTs are replacing the FETs and yielding roughly 15 K lower system temperatures at the zenith. The existing FET amplifiers will be replaced with HEMTs before the Neptune encounter.

The first 8.4 GHz amplifier (a FET) was installed in November 1984. There were two antennas with X-band capability until May 1986. Since then, X-band systems have been installed at the approximate rate of one every seven weeks. Through February 1987, there were seven antennas equipped with X-band capability, three with FET amplifiers and four with

HEMT amplifiers. As with the other VLA frequencies, each system consists of two intermediate frequency pairs. The right and left circularly polarized channels of one pair are designated "A" and "C," while those of the other pair are "B" and "D." Time on the array is allocated monthly in order to test system performance and operation with DSN telemetry receivers.

One of the tests that is often performed is a tip curve of all available X-band antennas in order to estimate the system temperatures including the atmospheric contribution. This is done for each antenna independently. Taken at face value, the tip curve data suggest unreasonably large variations of atmospheric temperature across the array, which implies inordinately large fluctuations of water vapor. We suspect that this result is due to two circumstances: (1) the *in situ* value of the calibration diode differs from its value determined in the laboratory; and (2) the actual value of the calibration diode varies slightly as a function of temperature and power supply current. We examine the nominal calibration of noise diodes on each antenna by assuming uniformity of the atmosphere and then derive a new value for the noise diode calibration level.

At various times, well-known radio sources are observed in order to estimate the antenna aperture efficiencies (i.e., the gain of the individual antennas). The ratio of the antenna gain to the system temperature, G/T , is an important parameter in the communications link performance and is derived for each antenna in operation by March 1987. In this article we summarize the results of the analysis of both the system temperature and aperture efficiency data from the past year.

II. System Temperature Data

The X-band configuration of the VLA antennas consists of dual amplifiers (one for each polarization) with a noise diode coupled to the input of each amplifier. The noise diodes can be switched on and off by one of the control computers and the output synchronously demodulated. Thus, the system is capable of being operated as a noise adding radiometer (NAR). The block diagram of the VLA electronics is shown in [1], and a description of the radiometric capability is given in [4]. The equivalent temperature of the noise diode is adjusted to be approximately 10 to 20 percent of the system temperature at the time each amplifier package is assembled and tested, but prior to installation on the antenna. This equivalent temperature of the noise diode is referred to as the "nominal value" and will vary slightly among the different amplifiers.

At X-band frequencies, the atmosphere radiates at an apparent temperature of several degrees; it can be used to

calibrate the noise diode equivalent temperature much like a cold load. The technique that is used is called a "tip curve" and has both advantages and disadvantages as compared with the laboratory determination. The primary advantage is that the tip curve is done *in situ* on the antenna. The disadvantages are the following: (1) an independent measurement of atmospheric water vapor is needed; (2) it is necessary to calculate the emission of both water vapor and oxygen with high accuracy (if an absolute calibration is desired); (3) the atmosphere must be assumed to be plane-layered; and (4) the system temperature is dominated by radiation in the main beam of the antenna. These factors will be discussed as the analysis proceeds.

VLA operations provide the capability of running a procedure called TIPPER that drives all antennas in a sub-array to elevation angles of 60, 40, 30, 25, 20, 15, and 10 degrees. NAR data is taken at each of these elevation angles as the antennas tip toward the horizon at an azimuth of roughly 90 degrees, and then again at each of these elevations as the antennas return to zenith. The procedure takes approximately 15 minutes and provides 13 estimates (the 10 degree point is not repeated) of the system temperature (i.e., the brightness temperature of the sky). The tip curve data that we will consider in this article are summarized in Table 1.

A gating circuit is used to measure the "gated total power" level when the noise diode is off, V_{gtp} . This includes all contributions to the system temperature except for the automatically switching noise diode. This gated total power level is kept very near 3 volts by an automatic level control. The quantity V_{sd} is the voltage from the "synchronous detector," and is the observable during a tip curve. It represents the extra voltage from a square-law detector that is synchronized with the noise diode switch rate of 9.6 Hz. That is, this voltage shows the level of contribution of the noise diode in proportion to the nominal system temperature. The synchronous detector has a gain of 15 relative to the gated total power detector. Hence, if the noise diode contributes 10 percent of the system temperature, the ratio V_{sd}/V_{gtp} is $15 \times 0.1 = 1.5$. To compensate for this additional gain, the voltage at the gated total power detector is multiplied by 15 in the equation for the system temperature. Therefore, we have the following relation for the system temperature observed when the noise diode is off:

$$T_{sys} \propto \frac{15V_{gtp}}{V_{sd}} \quad (1a)$$

A lower value of V_{sd} means that the noise diode contributes a smaller fraction of the system temperature, and hence that the overall system temperature is higher. Therefore, as an antenna

points to a lower elevation, the system temperature increases and V_{sd} decreases.

The nominal noise temperature contributed by the noise diode is T_{cal} , which is used to scale the voltage ratio in order to give the actual system temperature value. Hence,

$$T_{sys} \approx 15T_{cal} \frac{V_{gtp}}{V_{sd}} \quad (1b)$$

The real expression for the system temperature at a VLA antenna is slightly more complex, because it also includes measured DC offsets of the two voltage detectors, V_{gtp0} and V_{sd0} . Thus, the final estimate of system temperature is given by the expression [5]

$$T_{sys} = 15T_{cal} \frac{(V_{gtp} - V_{gtp0})}{(V_{sd} - V_{sd0})} \quad (1c)$$

Typically, V_{gtp} and the zero-level offsets, V_{gtp0} and V_{sd0} , change slowly and by very small fractional amounts. Operational procedure at the VLA is to measure these three values for each antenna intermediate frequency (IF) channel at roughly two-week intervals. Once determined, they are assumed constant and used in Eq. (1c) to calculate the system temperature. We will examine this assumption later, since changes in these values on time scales shorter than two weeks can lead to small errors in the system temperature or to apparent changes in the value of T_{cal} .

In the radiometry mode used for tip curves, an IF bandwidth of 50 MHz is used along with a minimum integration time of 1 s. Together with a T_{cal} value that is ~ 15 percent of T_{sys} , this gives an rms fluctuation level of approximately $10^{-3} \times T_{sys}$. That is comparable to the quantization level of the analog-to-digital converter, so that fluctuations occur in the least significant bit.

The system temperature looking at cold sky is

$$T_{sys} = T_{rec} + T_c e^{-\tau} + T_{atm} \quad (2)$$

Here, T_{rec} is the receiver temperature (in kelvins) referenced to the antenna aperture, T_c is the cosmic blackbody background temperature (2.8 K), T_{atm} is the temperature contribution by the atmosphere, and τ is the optical depth of the atmosphere in nepers. The atmospheric temperature is given by

$$T_{atm} = T_m (1 - e^{-\tau}) \quad (3)$$

where T_m is the mean radiating temperature of the atmosphere, in kelvins.

Note that the conventional definition of T_{rec} is referenced to the input of the first amplifier stage, whereas the definition made above includes losses and scattering in the antenna structure. If the atmosphere is plane layered then we can write,

$$\tau = \tau_0 \times AM \quad (4)$$

where τ_0 is the optical depth at the zenith (in nepers) and AM is the air mass ($AM \approx \csc$ [elevation angle]). Given a set of data at elevation angles E_1, E_2, \dots, E_N , the combination of Eqs. (1) and (2) gives a series of N equations that can be solved for some combination of the parameters T_{cal} , T_{rec} , and τ_0 .

In the first part of the analysis we will assume that the values for T_{cal} are accurate for each system and use the tip curve data to solve for τ_0 . Figure 1 shows the data from a typical tip curve with the system temperatures as calculated from Eq. (1) plotted versus the air mass. At the 8.4 GHz frequency the atmospheric opacity is typically small (i.e., $\tau \ll 1$), so the tip curve is linear with air mass. Note that the data are linear except for the point at 10 degree elevation (5.75 air masses).

The geometry of the feed and subreflector suggests that there is forward spillover that may be intercepting the land mask, which is typically greater than 0 degrees. At 10 degrees, this spillover picks up radiation from the earth at ~ 270 K and can contaminate the tip curve data even though the gain is reduced by 20 dB or more. For this reason we do not use any of the data from 10 degree elevation in the analysis. Figure 2 shows the residuals after we solved for the zenith opacity τ_0 . We plot the observed system temperature minus that calculated from the fit, versus the air mass. The data at 10 degrees were not used in the fit, but are shown in the plot to emphasize that they are anomalous on almost all of the antennas and are very repeatable on a given antenna. When all of the tip curve residuals are averaged (except the 10 degree points), the rms at each elevation is on the order of 0.1 K.

After we have solved for the zenith opacity we can calculate the temperature contribution of the atmosphere from Eq. (3) if we know the mean radiating temperature of the atmosphere, T_m . Using a nominal value $T_m = 257$ K, we analyzed the tip curve data listed in Table 1. Note that we could have included T_m as a solved-for parameter. However, as shown in the appendix, T_{atm} is relatively insensitive to this quantity. Figure 3(a-j) shows the solved-for atmospheric temperature versus the antenna number where the "AB" and "CD" notation is the convention at the VLA to designate the two circular polarizations, right and left, respectively. Note

that antenna 10 came on line in February 1987 and is missing from the earlier tip curve data, and antenna 24 is missing from the 24 and 25 November data due to equipment problems.

Consider two aspects of these plots: first, the variation in the average level from one figure to the next; and second, the shape of each plot. Also note that although the average atmospheric temperature varies from day to day as we might expect, the general shape of the curve is remarkably consistent from month to month. The antenna-to-antenna variation is larger than we would expect from atmospheric fluctuations over the baselines in the array. The common element in this part of the analysis is the fact that we used the nominal values for T_{cal} . In the case of a perfectly stable, calibrated system and a stratified atmosphere, we would expect that the plots in Fig. 3 would be nearly flat; i.e., all antennas would report nearly the same value for T_{atm} . In reality, some variation would be expected due to (1) real variations in atmospheric water vapor or liquid (e.g., clouds or rain cells); (2) system noise; (3) changes in the offset calibration values; or (4) changes in the value of the noise diode calibration T_{cal} .

We might expect the average value of T_{atm} to vary from day to day due to slow changes in the atmosphere, changes dominated primarily by the water vapor content and liquid water in clouds. It is a relatively straightforward matter to solve the equation of radiative transfer and compute the apparent brightness temperature of the atmosphere at the frequency 8.4 GHz, given the amount and distribution of water vapor and oxygen. It is found that it takes 0.42 cm of precipitable water vapor to cause a 0.1 K change in T_{atm} at this frequency. This amount of water vapor is equivalent to a difference in propagation delay of 2.6 cm at the zenith over distances of a few kilometers. For a typical wind speed of 8 m/s, this would imply phase changes of three-fourths of a turn on time scales less than 500 sec. It is highly unlikely that the vapor over the array will differ by this amount, so we conclude that the assumption of constant atmosphere over the array is a good one.

As mentioned above, small changes in the offset data can cause systematic errors in the estimated brightness temperatures, or in our case changes in the derived values of T_{atm} . In order to estimate the size of these effects we used the tip curve data from the first tip curve on 26 October 1986 and reduced it using the offset data from 24 October 1986, 24 November 1986, 16 December 1986, 14 January 1987, and 19 February 1987. Table 2 summarizes the average as well as maximum positive and negative deviations from the first set of offset data. We see that although the average changes are quite small, there are "jumps" in the offset data that can cause relatively large changes in the solved-for parameters.

The offset data for antennas 11AB and 11CD are listed in Table 3. The change in V_{sd0} between 24 October and 24 November 1986 gives rise to changes of -0.131 K, -0.45 K, and -0.00052 neper in T_{atm} , T_{rec} , and τ_0 , respectively. Similarly, the change in V_{gtp} between 14 January and 19 February 1987 causes changes of 0.144 K, 1.43 K, and 0.00057 neper in T_{atm} , T_{rec} , and τ_0 . We conclude that, although the average values of these quantities are quite stable, variations on time scales of approximately two weeks or less in the offset data can lead to variations of ± 0.15 K in the atmospheric temperature T_{atm} , ± 1.4 K in the receiver temperature T_{rec} , and ± 0.0006 in the zenith opacity τ_0 for a particular observation. If higher accuracy is required, calibration data should be sampled both before and after an observation. It seems unlikely that the offset data variations could give the repeatable signature shown in Fig. 3 over a six-month period.

The system noise can be estimated from the repeatability of the two tip curves taken on 26 October, about 1 hour apart, and comparing the solved-for parameters from each antenna channel. The average changes as well as maximum increases and decreases from the first to the second tip curve are listed in Table 4. We note that the absolute surface humidity values, as indicated by the temperature and dew point measurements, were 3.6 and 3.5 gm/m³, respectively, for the start of the two tip curves, suggesting a stable atmosphere during the observations. We conclude that system noise can account for some of the small differences in Fig. 3(a-j) but cannot account for the large differences between antennas or for the repeatability of the signature.

If variations in the atmosphere, DC offset data, and system noise cannot account for the differences seen in Fig. 3(a-j), we are left with the possibility of variations in the calibration of the noise diode. The precise value of T_{cal} is determined in the laboratory for each individual amplifier by comparison with a cold load, using the technique described in [6]. The laboratory calibration of the diode equivalent temperature involves some intrinsic measurement uncertainty. This calibration may be different from the actual value on the antenna due to the slightly different impedances of the cold load and the antenna as seen by the amplifier. If we use the laboratory determination of T_{cal} , this would lead to a bias in the determination of system temperature. In addition, the noise diode is sensitive to the physical temperature of its surroundings (typically 0.01 dB per kelvin) and to small changes in the current through the device. Since the front-end area on the antenna is reasonably stable (± 2 K), we might expect T_{cal} to change slowly around some nominal value due to temperature cycling. The temperature cycling is impossible to solve for without knowing the physical temperature variations in the front-end area. However, a potential bias in T_{cal} for each antenna can be extracted, at least in a relative sense.

In the second part of the analysis we adopt as the reference antenna 11AB. In this analysis, we force all antennas to “see” the same opacity as 11AB and solve for the value of T_{cal} for each tip curve. The 10 values for T_{cal} are then averaged to estimate the possible bias relative to antenna 11AB. Table 5 summarizes the nominal values of the noise diode calibration from the laboratory determination versus the solved-for values determined by the above procedure.

If we now use these new values of T_{cal} to reanalyze the tip curves, the variations in T_{atm} are much less pronounced. The dotted curves in Fig. 3(a-j) show the solved-for atmospheric temperature using these corrected T_{cal} values, and Table 6 summarizes the rms values obtained from using the nominal values of T_{cal} versus the solved-for values of T_{cal} . We see that the rms of T_{atm} is reduced by approximately a factor of two in this procedure and conclude that these new values of T_{cal} are more internally consistent than the nominal values. We have probably removed most of the bias term in T_{cal} that is due to differences in the laboratory value versus its *in situ* value. We have not reduced variations in T_{cal} around its average value, nor have we addressed its absolute value. Table 7 lists the zenith values for T_{sys} and T_{rec} derived using the new values of T_{cal} . Also, note that although the variation in T_{atm} from antenna to antenna is reduced, the rms is still more than we might reasonably expect from fluctuations in water vapor over the array.

III. Aperture Efficiency Measurements

No antenna does a perfect job of detecting the radiation that impinges on its primary surface. Some of the effects that reduce the efficiency of an antenna are diffraction effects, physical blockage by support structures, irregularities in the main and the subreflector shape, imperfect illumination, mechanical alignment of the feed and subreflector, subreflector focus, and pointing. The aperture efficiency is a measurement of the fraction of the radiation hitting the main reflecting surface that is actually detected; it is the ratio of the “effective” area of an antenna to its actual physical area. For monolithic radio telescopes operating at their primary observing frequencies, aperture efficiencies are typically 50–70 percent. The efficiency also tends to be reduced when observations are made at low elevation angles because of increased deformation of the telescope surfaces.

The VLA antennas were designed to work well at frequencies up to 22 GHz. Therefore, good aperture efficiencies should be expected at 8.4 GHz if the new X-band systems are designed and installed properly. Measurements of the aperture efficiencies of the antennas with X-band systems thus serve to verify the overall design and ensure that the hardware is installed properly on the telescopes.

Aperture efficiencies are typically found by measuring the antenna temperatures given by observations of radio sources with known flux densities. The effective area of a telescope is given by the formula $A_e = 2kT_{\text{ant}}/S$, where k is Boltzmann’s constant, T_{ant} is the antenna temperature contributed by the radio source, and S is the radio source flux density. Since the effective area is simply the aperture efficiency ϵ multiplied by the physical area A_p , the efficiency is then given by the equation,

$$\epsilon = \frac{2kT_{\text{ant}}}{SA_p} \quad (5)$$

Frequently, this expression is rearranged to give an expression for the overall sensitivity of a telescope,

$$\frac{T_{\text{ant}}}{S} = \frac{\epsilon A_p}{2k} = \frac{A_e}{2k} \quad (6)$$

In normal radio astronomy usage, this sensitivity is expressed in kelvins per jansky, where the jansky is the unit of radio flux density and is defined to be equal to 10^{-26} W/m²/Hz.

The antenna temperature is simply the difference between the system temperature measurements on and off source. The system temperature of a VLA antenna is measured using the NAR and calculated using the expression for T_{sys} given in the previous section. Therefore, the final expression for the aperture efficiency (dimensionless) of a VLA antenna becomes:

$$\epsilon = \frac{30kT_{\text{cal}}(V_{gtp} - V_{gtp_0})}{SA_p} \left[\frac{1}{(V_{sd} - V_{sd_0})_{\text{on}}} - \frac{1}{(V_{sd} - V_{sd_0})_{\text{off}}} \right] \quad (7)$$

The VLA has a standard observing mode which can be used to measure aperture efficiencies. This mode is somewhat more elaborate than just making simple measurements on and off source. Instead, the antennas point at a specified source position and at four points separated from that position by a distance equal to the nominal half-power half-width of the beam. Two points are offset by plus or minus half a beamwidth in elevation, while the other two are offset by plus or minus half a beamwidth in azimuth. Interleaved among the “on-source” and “half-power-point” measurements are observations that are made 5 beamwidths off source in azimuth. At the 8.4 GHz operating frequency the antenna has a half-power beamwidth of 90 millidegrees and the peak pointing residuals are typically 2 to 3 millidegrees, so the effective pointing loss is less than 1 percent. However, there is a known beam squint between the two polarizations that causes point-

ing offsets of 5 millidegrees in azimuth and 1.5 millidegrees in elevation at 8.4 GHz. We have chosen to optimize the pointing for both polarizations simultaneously so as to use a single set of pointing offsets. This results in measured antenna efficiencies that are systematically 1–2 percent lower than theoretically possible.

The procedure for determining the aperture efficiency via measurements of a given source is straightforward. The fringe amplitude on a strong point source is measured relative to the standard flux density calibrator 3C286, whose flux density is assumed to be 5.20 Jy at X band. This gives the value for S to be used in the efficiency formula. The observing mode described above is used to measure the synchronous detector voltage both on and off source. If necessary, the measurements made at the half-power points are used to correct the on-source measurements for possible pointing errors. Such corrections are not made if the pointing has been checked prior to the observing session and found to be adequate. Then, the measured numbers are used to calculate the efficiency for each polarization of each available antenna. When time permits, measurements on several different sources are made consecutively.

Aperture efficiency measurements have been made periodically since the beginning of 1986. Since only two X-band antennas were available until about mid-1986, the histories on other antennas cover no more than nine months. Table 8 summarizes the results for each antenna measured to date, using the solved-for values of T_{cal} . Elevation-angle dependence is ignored in this table. Right and left circularly polarized data are distinguished, and the rms error is also given. Quoted errors are statistical only, and do not take possible systematic effects into account. One IF pair for antenna 11 (the “BD” pair) is not included, because it consistently gives abnormally high efficiency values, presumably due to some instrumental problem that has not yet been tracked down. Values for antenna 10 should be considered preliminary, since it was recently brought on line and was measured only via observations of three different radio sources in February 1987. Note that the efficiency for antenna 20 is considerably lower than that for the other antennas. Antenna 20 was the first VLA antenna to be equipped at X band; it uses a nonstandard feed and receiver rather than the production models that were installed on the other antennas. The preliminary feed design is thought to account for the low efficiency and will be swapped for a production model prior to the encounter.

Summaries for various combinations of data are given at the bottom of Table 8. For reasons stated above, no data from antenna 20 are included in the summaries. The agreement among the other antennas is fairly good, although there are

some discrepancies such as the apparently low efficiency for the left circularly polarized feed of antenna 24.

The three radio sources used most often for aperture efficiency measurements were 3C273, 3C279, and 3C345. In general, 3C273 gave higher aperture efficiencies than the other two sources. This may be caused by the fact that 3C273 is about three times stronger at X band than are the other two sources. It contributes slightly more than 3 K, or about 10 percent for a HEMT-equipped antenna, to the total system temperature. The other two sources contribute only about 3 percent to the system temperature. Detector nonlinearities or differing signal-to-noise ratios may account for the apparent difference in aperture efficiency. No other sources have been measured often enough to confirm that the possible effect is actually dependent on source strength. Since 3C273 and 3C279 are much further south than 3C345, they were typically measured at lower elevations. The data listed in Table 8 were taken at elevation angles between 30 and 45 degrees and are not sufficient to determine potential elevation angle dependence.

Including the possible systematic errors caused by noise diode variations and by measurements of different sources, the best estimate for the aperture efficiencies of the VLA X-band systems is 0.62 ± 0.03 at elevations of ~ 40 degrees. The antenna sensitivities are then computed to be 0.110 ± 0.005 K/Jy.

IV. Ratios of Gain to System Temperature

The overall figure of merit of an antenna is determined by a combination of its noise level and its gain. For a VLA antenna at X band, the ratio of gain to system temperature (hereafter G/T) is given by

$$G/T = 4.859 \times 10^6 \frac{\epsilon}{T_{\text{sys}}} \quad (8)$$

The numerical factor in this equation is derived from the definition of the gain, $G = 4\pi A_e / \lambda^2$ (e.g., [7]). The symbol λ represents the observing wavelength.

Note that the value of G/T depends only on the aperture efficiency and system temperature of an antenna of a given physical size. The preceding sections have described the separate calculations of the aperture efficiencies and system temperatures. One of the major uncertainties in the determination of each quantity individually is the possible error or variation in the effective temperature of the noise diode, T_{cal} . However, as shown by Eqs. (1) and (7), both aperture efficiency and system temperature values are directly proportional to the assumed temperature of the noise diode. There-

fore, in the evaluation of Eq. (8), the assumed value of T_{cal} divides out and is removed as a source of error in the determination of antenna figure of merit.

The values of G/T for all of the VLA antennas equipped at X band are shown in Table 9. For four HEMT-equipped antennas (numbers 3, 10, 11, and 24), the weighted mean value of G/T is found to be $9.40 \times 10^4 \text{ K}^{-1}$ at the zenith, with an rms of $6.0 \times 10^3 \text{ K}^{-1}$. The value of G/T at 30 degree elevation is also of importance; since the Voyager spacecraft is at a southerly declination, its maximum elevation at the VLA during the Neptune encounter phase will be ~ 34 degrees. Assuming an atmospheric temperature of 2.7 K per air mass, we find $G/T = 8.65 \times 10^4 \text{ K}^{-1}$ at 30 degree elevation. By comparison, we note that the most sensitive VLA frequency prior to the X-band installation was 5.0 GHz (C band). At that band, the VLA antennas have typical system temperatures of 50 K and aperture efficiencies of 0.65 at zenith. Using these values and the wavelength of X band, we calculate $G/T = 6.32 \times 10^4 \text{ K}^{-1}$ at zenith. Thus, the VLA antennas have a typical G/T at X band that is about 50 percent higher than would be expected if their overall performance were the same as at C band. This is consistent with the fact that the antennas that were initially equipped with FETs at X band give measured values of G/T that are near $6 \times 10^4 \text{ K}^{-1}$.

At 30 degree elevation, the sum of 27 VLA antennas (assuming all HEMT amplifiers) with zero relative phase would give a net G/T of $2.34 \times 10^6 \text{ K}^{-1}$, or 63.7 dB/K. Including a 1.0 dB loss because of the 3-level quantization and

the periodic data gap at the VLA, the final estimate of G/T is then 62.7 dB/K. By comparison, a 64-m antenna of the Deep Space Network (DSN), assumed here to have a 50 percent aperture efficiency and a system temperature of 25 K at 30 degree elevation, would have $G/T = 58.0 \text{ dB/K}$. Therefore, the performance of the individual antennas at the VLA predicts that the array of 27 antennas will give an overall enhancement of 4.7 dB, or a factor of 2.9, over the G/T of a single 64-m DSN antenna. There will likely be a modest signal loss of 0.2 to 0.3 dB because the 27 VLA antennas will have non-zero relative phases caused by system noise and by imperfectly corrected tropospheric disturbances. Thus the net improvement afforded by the VLA will be about 4.4 dB relative to a single 64-m antenna, making the VLA "worth" approximately 2.75 64-m antennas.

V. Summary

The system temperatures, aperture efficiencies, and gain to system temperature ratio for the 7 VLA antennas that were equipped with X-band amplifiers as of March 1987 have been measured. An internally consistent set of system temperature calibrations was derived referenced to antenna 11 and listed in Table 7. Aperture efficiencies, as summarized in Table 8, were derived by observing several strong radio sources whose flux densities were measured relative to the calibration source 3C286. The G/T ratio for the HEMT-equipped antennas was found to be $9.40 \times 10^4 \text{ K}^{-1}$ at the zenith. The fully phased VLA is predicted to have a G/T value that is 4.4 dB above that of a single 64-m DSN antenna.

References

- [1] P. J. Napier, A. R. Thompson, and R. D. Ekers, "The Very Large Array: Design and Performance of a Modern Synthesis Radio Telescope," *Proc. IEEE*, vol. 71, pp. 1295-1320, 1983.
- [2] J. W. Layland and D. W. Brown, "Planning for the VLA/DSN Arrayed Support to the Voyager at Neptune," *TDA Progress Report 42-82*, vol. April-June 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 125-135, August 15, 1985.
- [3] S. M. Petty, "Microwave Devices," in *Low-Temperature Devices*, R. K. Kirshman (ed.), New York: IEEE Press, 1986, pp. 358-363.
- [4] A. R. Thompson, *An Introduction to the VLA Electronics System*, VLA Technical Report No. 29, Charlottesville, Virginia: NRAO, March 1977.
- [5] "GTTSYS—System Temperature Entry Into Gain Table," in *VLA Observers' Reference Manual*, Socorro, New Mexico: NRAO, 1987.
- [6] S. Weinreb, *Calibration of HP-346B Noise Sources at 1.3-1.7 GHz*, NRAO Electronics Division Technical Note No. 111, Charlottesville, Virginia: NRAO, July 1982.
- [7] J. D. Kraus, *Radio Astronomy*, New York: McGraw-Hill, pp. 212-214, 1966.
- [8] G. M. Resch, "Inversion Algorithms for Water Vapor Radiometers Operating at 20.7 and 31.4 GHz," *TDA Progress Report 42-76*, vol. October-December 1983, Jet Propulsion Laboratory, Pasadena, California, pp. 12-26, February 15, 1984.

Table 1. Summary of tip curve observations*

Date	Time	Temperature	Dew point	Remarks
26 Oct. 1986	1000	9.0	-3.9	Clear and dry
26 Oct. 1986	1050	11.9	-4.6	Clear and dry
27 Oct. 1986	1340	16.3	-5.5	Clear and dry
24 Nov. 1986	1110	2.7	-9.2	Clear
25 Nov. 1986	1350	9.8	-4.1	Clear
16 Dec. 1986	1030	3.2	1.7	Heavy overcast, drizzle
13 Jan. 1987	0800	-9.0	-10.9	Clear
19 Feb. 1987	1200	-3.9	-7.3	Cloudy, fog, snow on dishes
27 Feb. 1987	0140	-4.4	-7.5	Clear and calm
1 Mar. 1987	0140	-8.0	-9.4	Clear and calm

*Time is the start in mountain standard time zone; surface temperature and dew point near the array center are given in degrees Celsius.

Table 2. Changes in solved-for parameters due to changes in the offset data

Parameter	Average change	Maximum increase	Maximum decrease
T_{atm} (K)	0.0007	0.144	-0.131
T_{rec} (K)	0.09	1.43	-0.45
Opacity (nepers)	$<10^{-5}$	0.00057	-0.00052

Table 3. Summary of offset data (in volts) for antennas 11AB and 11CD

11AB				11CD			
Date	V_{gtp_0}	V_{gtp}	V_{sd_0}	Date	V_{gtp_0}	V_{gtp}	V_{sd_0}
24 Oct. 1986	-0.08	3.01	-0.094	24 Oct. 1986	-0.03	3.02	0.005
13 Nov. 1986	-0.08	3.00	0.041	13 Nov. 1986	-0.03	3.02	0.004
10 Dec. 1986	-0.04	2.99	0.074	10 Dec. 1986	-0.03	3.02	0.005
14 Jan. 1987	-0.03	3.02	-0.015	14 Jan. 1987	-0.04	3.06	-0.004
19 Feb. 1987	-0.03	3.02	0.018	19 Feb. 1987	-0.04	2.85	0.008

Table 4. Changes of solved-for parameters from successive tip curves obtained on 26 October 1986

Parameter	Average change	Maximum increase	Maximum decrease
T_{sys} (K)	-0.13	0.06	-0.27
T_{atm} (K)	-0.0084	0.062	-0.082
T_{rec} (K)	-0.12	0.14	-0.33
τ_0 (neper)	0.00003	0.00025	-0.0003

Table 5. Nominal and solved-for values of the noise diode value T_{cal} (in kelvins)*

Antenna	T_{cal}		rms	Ratio [†]
	Nominal	Solved-for		
3AB	4.20	4.71	0.24	0.892
3CD	4.33	4.70	0.26	0.922
10AB	4.33	4.68	0.08	0.925
10CD	4.50	4.59	0.10	0.980
11AB	4.03	4.03	0.001	0.9999
11CD	3.94	3.92	0.17	1.004
20AB	5.00	4.91	0.21	1.019
20CD	4.20	4.69	0.22	0.896
21AB	3.70	3.70	0.16	1.001
21CD	4.00	3.85	0.14	1.039
24AB	4.32	4.48	0.17	0.965
24CD	4.21	4.02	0.14	1.047
25AB	4.07	4.22	0.25	0.965
25CD	3.99	4.17	0.28	0.957

*Antenna 11AB was used as the reference for the opacity, and T_{cal} was computed on all other channels to give the same opacity for each tip curve.

[†]Ratios are those of the nominal values to the solved-for values of the noise diode temperatures.

Table 6. Comparison of the atmospheric temperature (in kelvins) derived from tip curve data, using two different values of T_{cal} for each antenna IF*

Date	Before correction		After correction	
	T_{atm}	rms	T_{atm}	rms
Oct. 26 (a)	2.66	0.13	2.73	0.05
Oct. 26 (b)	2.67	0.11	2.74	0.05
Oct. 27	2.70	0.12	2.77	0.06
Nov. 24	2.58	0.22	2.64	0.13
Nov. 25	2.57	0.22	2.65	0.14
Dec. 16	2.81	0.24	2.90	0.16
Jan. 13	2.59	0.13	2.67	0.04
Feb. 19	2.89	0.16	2.97	0.14
Feb. 27	2.62	0.13	2.70	0.06
Mar. 1	2.57	0.13	2.64	0.05

*Nominal values of T_{cal} were used first and then corrected by taking antenna 11AB as a reference.

Table 7. Average zenith system temperature and receiver temperature (K) for each antenna, as found using the solved-for values of T_{cal}

Antenna	T_{sys}	rms	T_{rec}	rms
3AB	33.1	1.4	27.6	1.4
3CD	34.8	1.2	29.3	1.3
10AB	37.5	1.4	32.0	1.3
10CD	34.8	1.2	29.2	1.1
11AB	31.1	0.5	25.6	0.4
11CD	29.9	0.8	24.4	0.7
20AB	44.7	0.5	39.2	0.5
20CD	49.5	0.5	44.0	0.4
21AB	46.5	0.6	41.0	0.6
21CD	47.8	0.8	42.3	0.7
24AB	27.5	1.0	22.0	1.0
24CD	29.4	1.0	23.9	1.0
25AB	51.7	1.2	46.1	1.2
25CD	51.7	1.2	46.2	1.2

Table 8. Aperture efficiency measurements of VLA antennas

Antenna	Receiver	RCP		LCP		Total	
		ϵ	rms	ϵ	rms	ϵ	rms
3	HEMT	0.659	0.061	0.661	0.046	0.660	0.052
10	HEMT	0.636	0.031	0.631	0.033	0.634	0.032
11AC	HEMT	0.621	0.029	0.606	0.031	0.614	0.030
20	FET	0.545	0.047	0.570	0.036	0.561	0.040
21	FET	0.603	0.048	0.599	0.044	0.601	0.046
24	HEMT	0.614	0.028	0.567	0.042	0.600	0.033
25	FET	0.614	0.061	0.619	0.053	0.617	0.057
Weighted mean (no 20)		0.623	0.037	0.613	0.039	0.618	0.038
Weighted mean (HEMT)		0.625	0.033	0.615	0.037	0.621	0.035
Weighted mean (3C273)		0.651	0.041	0.637	0.033	0.643	0.036
Weighted mean (3C345)		0.607	0.036	0.591	0.042	0.600	0.039
Weighted mean (3C279)		0.620	0.045	0.617	0.049	0.619	0.047

Average source fluxes:
 $S(3C273) = 31$ Jy
 $S(3C345) = 11$ Jy
 $S(3C279) = 10$ Jy

Table 9. List of G/T values (units of $10^4/K$)

Antenna	Receiver	RCP		LCP		Total	
		G/T	rms	G/T	rms	G/T	rms
3	HEMT	9.67	0.98	9.23	0.72	9.38	0.82
10	HEMT	8.24	0.51	8.81	0.55	8.50	0.53
11AC	HEMT	9.70	0.48	9.85	0.57	9.76	0.52
20	FET	5.92	0.51	5.60	0.36	5.71	0.42
21	FET	6.30	0.51	6.09	0.46	6.18	0.48
24	HEMT	10.85	0.63	9.37	0.76	10.25	0.69
25	FET	5.77	0.59	5.82	0.52	5.80	0.55
Weighted mean (HEMT)		9.47	0.58	9.31	0.63	9.40	0.60
Weighted mean (FET, no 20)		6.07	0.55	5.97	0.49	6.01	0.52

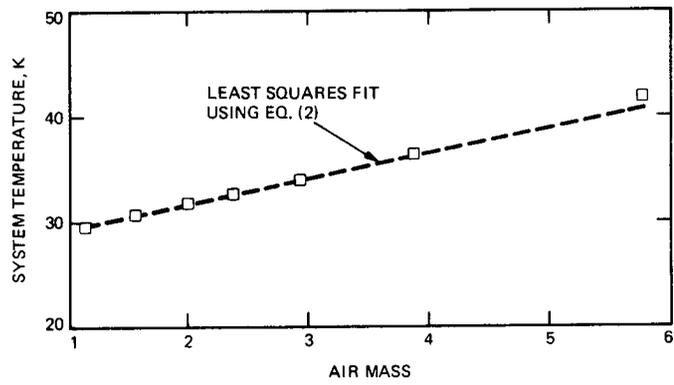


Fig. 1. Data plotted from antenna 3, channel AB, showing the total system temperature versus the air mass, illustrating a typical tip curve. Note that the datum at 10 degree elevation ($AM = 5.8$) is plotted but not used in the fitting procedure.

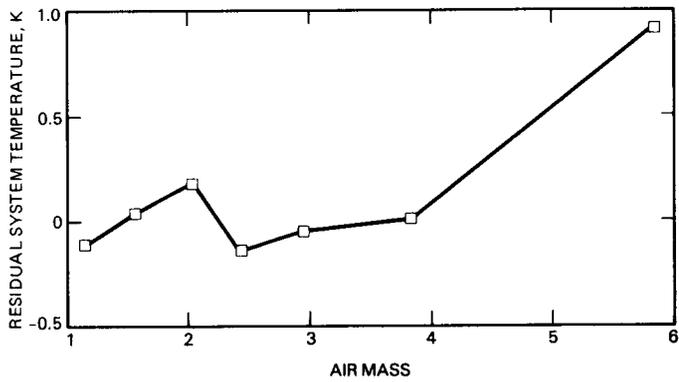


Fig. 2. Plot of the residuals from Fig. 1. Neglecting the point at $AM = 5.8$, the residuals are typically less than 0.2 K.

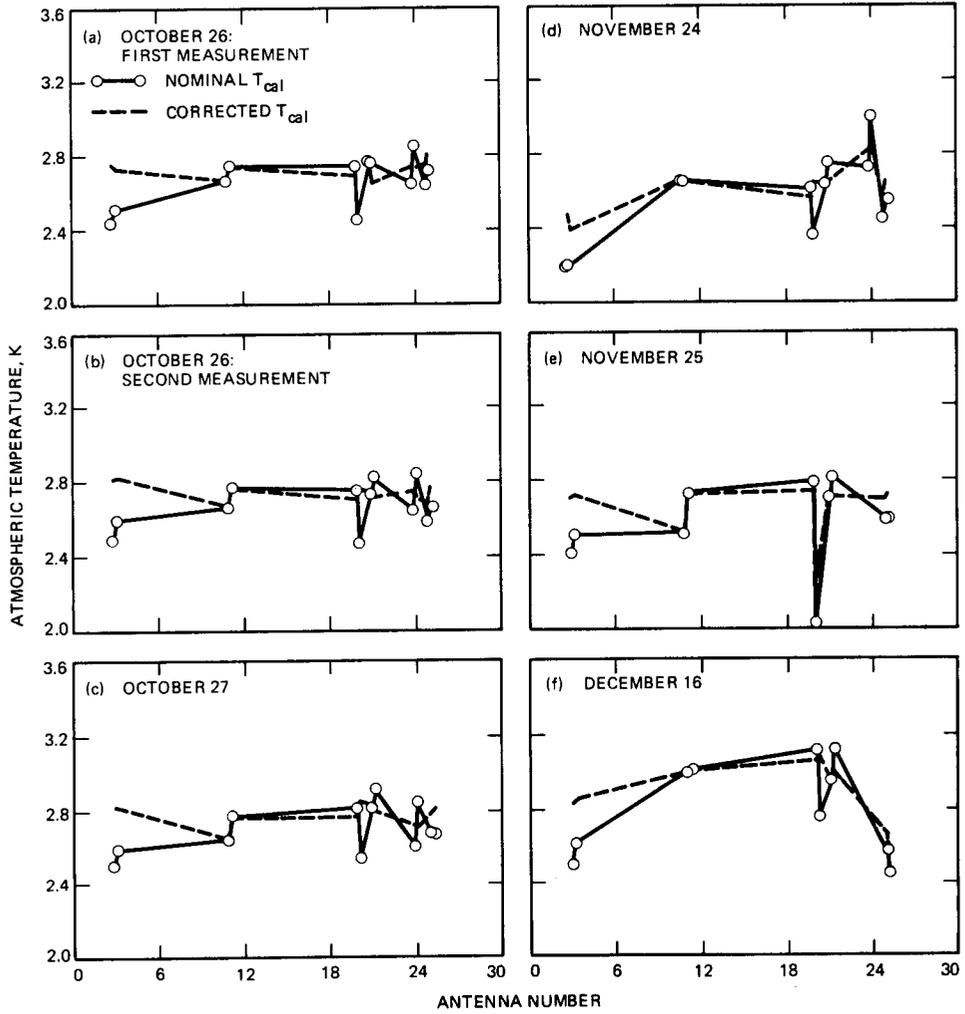


Fig. 3. For each of the tip curves listed in Table 1, these plots show the atmospheric temperature solved for from Eqs. (1) and (2), versus the antenna number. The solid line connects the data points indicating the solution using the nominal values of T_{cal} . The dotted lines indicate the values of T_{atm} using the solved-for values of T_{cal} listed in Table 5. Note the reduced variation in the dotted lines when compared to the solid lines.

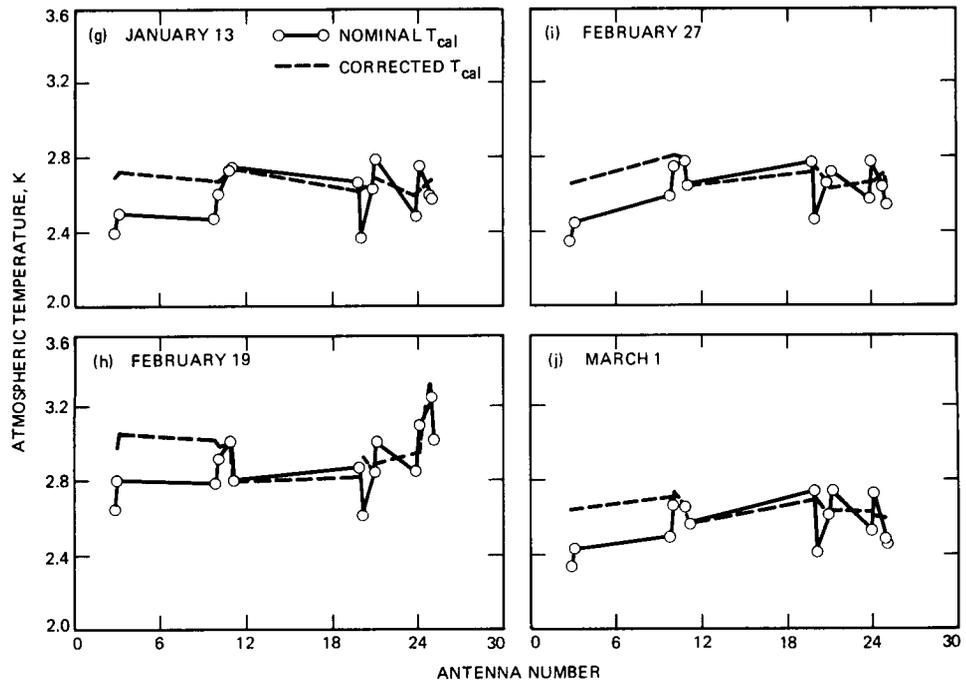


Fig. 3 (contd)

Appendix

Calculation of T_m

An accurate calculation of T_m , the mean radiating temperature of the atmosphere, depends not only on the vertical temperature distribution of the atmosphere at a given time but also on the fact that the two primary molecular contributors (water vapor and oxygen) have different vertical density distributions. Some of the variations observed in T_m from sites within the continental United States are discussed in [8]. For the present purpose we have made use of one year (1976) of radiosonde data from El Paso, Texas, in order to estimate T_m from surface temperature data. El Paso is approximately 280 km from the VLA and is in approximately the same climatic zone, although El Paso is at an altitude of 1200 m versus 2124 m for the VLA site. Using the El Paso data and assuming a linear relation between T_m and the surface temperature T_s gives

$$T_m = 256.9 + 0.445T_s \quad (\text{A-1})$$

where T_s is in degrees Celsius. Note that the sensitivity of T_{atm} with respect to changes in T_m for $\tau_0 \ll 1$ is

$$\frac{dT_{\text{atm}}}{dT_m} = \tau_0 \quad (\text{A-2})$$

or

$$\frac{dT_{\text{atm}}}{dT_s} = 0.445 \tau_0 \quad (\text{A-3})$$

For the interval spanned by the data in this article, the zenith opacities are on the order of 0.01 and the surface temperatures vary by roughly 25°C. Hence, we would expect less than 0.1 K variation in T_{atm} from variations in T_m .

08

DSS 14 64-Meter Antenna S- and X-Band Efficiency and System Noise Temperature Calibrations, September 1987

S. D. Slobin

Radio Frequency and Microwave Subsystems Section

This article is the third in a series documenting the efficiency and noise temperature characteristics of the DSN 64-meter antenna network prior to its upgrading to 70-meter configuration. DSS 14 (Goldstone, California) is the last of the three large antennas to be upgraded, and the test results presented here document its performance just prior to its downtime during the end of 1987. Antenna area efficiency was found to be somewhat higher at DSS 14 than at DSS 43 (Australia) and DSS 63 (Spain). The peak X-band efficiency was determined to be 49.8 percent (without atmosphere), compared with 45.4 percent and 45.1 percent for DSS 43 and DSS 63, respectively. The X-band zenith system noise temperature was found to be 1 to 3 kelvins higher than at the other two stations, depending on which maser was chosen for the measurements. Ascribing efficiency differences to small-scale antenna surface roughness, DSS 14 may be regarded as having a 1.5- to 1.6-mm rms surface as compared to the other two antennas with 1.7- to 1.8-mm rms surfaces.

I. Introduction

This article is the third in a series documenting the performance of the DSN 64-meter network prior to its upgrading to 70-meter configuration. DSS 14 is the third of the three large antennas to be modified; DSS 63 was upgraded in May 1987, and DSS 43 in September 1987. DSS 14 will achieve its final configuration by early 1988. The S- and X-band performance of the overseas 64-meter antennas is documented in [1] and [2].

Antenna calibration measurements were taken during the months of August and September 1987. Because of a change

in radiometer calibration technique in the middle of the calibration project (hourly noise diode calibration instead of approximately once every 6 hours), only those data from days 258, 267, and 270 were used in the final data reduction. The radio sources used on those days as standard efficiency calibrators were 3C123, 3C274, and 3C286. Subsequent to the DSS 14 tests, modifications were made in the DSN radio source list¹ which affected these sources. The changes involved both flux and source size corrections, and it was necessary to

¹M. Klein, A. Freiley, and P. Richter, *DSN Radio Source List for Antenna Calibration*, JPL Report D-3801, Rev. B (internal document), Jet Propulsion Laboratory, Pasadena, California.

apply these corrections to the measured efficiency values after the fact. Since 3C274 is one of the strongest DSN calibrators visible from both the northern and southern hemispheres (and is also used in both the DSS 43 and DSS 63 calibrations), it must be accepted that the efficiency calibrations based on this source, as published in [1] and [2], are slightly in error. The source size correction for 3C274 increased 0.369 percent. This would have the effect of increasing the apparent antenna efficiency by approximately half this amount (for a 50 percent efficient antenna). The effect of these changes on the published DSS 43 and DSS 63 performance values will be discussed later.

The DSS 63 calibration article [1] contains a great amount of detail regarding calibration methods used and techniques involved in data reduction and analysis. That reference should be used if any uncertainty of meaning arises in the reading of this article.

Data taken on the three listed days appear to be generally well-behaved, and the weather was noted as clear with wind not exceeding 15 mph. Psychrometric data were taken hourly during the measurements, and this proved to be extremely useful in determining adjustments for the “no-atmosphere” antenna calibration values. Based on the temperature and relative humidity values given in the calibration report, an average weather model for DSS 14 was developed for the 13-day span of measurements. Total atmospheric attenuation was determined to be surprisingly constant over this time period, even though the temperature and relative humidity values varied greatly (e.g., the temperature ranged from 58°F to 103°F, and the relative humidity varied, almost inversely, from 89 percent down to 7 percent). For the purposes of this article, the S-band zenith attenuation was determined to be 0.025 dB, and the X-band zenith attenuation was determined to be 0.033 dB. These may be contrasted with the DSS 43/63 model of 0.03 dB and 0.04 dB, respectively, for S- and X-band, for year-average attenuation at those temperate locales.

After the data were corrected for flux and source size as given in [3], it became clear that the efficiency values determined using 3C123 were about 3 percent high (about 1-1/2 efficiency percent) when compared to the efficiency values determined using sources 3C274 and 3C286. This difference was the same at S- and X-bands, and did not seem to be a function of elevation angle (which would rule out a weather model error for data taken on different days). After consultation (M. J. Klein, private communication), it was decided to adjust the S- and X-band efficiency values by the factor 0.97 for those antenna efficiencies determined using radio source calibrator 3C123. This radio source calibrator is about one-fourth as strong as the strongest DSN standard calibrator (3C274), and since 3C274 and 3C286 give results agreeing very

well with one another, 3C123 efficiency determinations were changed. It is possible that 3C123 flux has changed; however, at this time, the source of the discrepancy has not been determined. The problem will be examined in the future, and 3C274/3C123 comparisons using other antennas at Goldstone may help resolve it. Although both 3C274 and 3C123 were used in the DSS 63 calibration [1], the 3 percent difference was not apparent in those data.

II. Antenna Area Efficiency

Figures 1 and 2 show the S-band antenna area efficiencies both with and without the atmospheric attenuation included. Note that area efficiency is referenced to a uniformly illuminated aperture, 64 meters in diameter, at the given frequencies. For S-band (2295 MHz), the 100 percent efficient antenna gain is 63.75 dBi; for X-band (8420 MHz) it is 75.04 dBi.

It should be noted that data taken at the low elevation angles (less than 10 degrees) show much scatter. This is possibly due to atmospheric attenuation changes from point to point or scintillations that affect the noise temperature measurement in the on-source position. Also shown in the figures are second-order curve fits to the data. Note that the curve fit in Fig. 2 (S-band efficiency without atmosphere) actually curves slightly upward! This is obviously a curve-fitting anomaly, with a possible contribution from a slight atmospheric model error. It would seem reasonable to assume that the S-band efficiency is nearly a constant 59.4 percent at all elevation angles. The data do not warrant a more complex description.

Figures 3 and 4 show the X-band area efficiencies both with and without atmosphere. Note again the extreme spread of measured values at elevation angles below 30 degrees. This undoubtedly contributes to the large uncertainties in the shape of the curve, even though the tight clustering of points in the 40- to 70-degree elevation region appears to determine the peak value fairly well.

Table 1 gives the coefficients of the second-order curve fits in Figs. 1 through 4. Also given in the table are peak values of efficiency and the elevation angles at which they occur.

III. System Noise Temperature

Figures 5 and 6 show the S-band system noise temperature both with and without atmospheric contribution. The two sets of data represent data taken with two different masers. The upper curve shows data taken with the Block V maser; the lower curve is that taken with the SPD maser in a low noise path. Because of the limited elevation angle range of the lower data set, it was not curve-fitted to represent system noise tem-

perature as a function of elevation angle. The upper data set in Figs. 5 and 6 was fitted with a fourth-order curve, the coefficients of which are given in Table 2. Note that in Fig. 5 (above 65 degrees) and in Fig. 6 (above 70 degrees) the extrapolated values of the curve fit are created to be constant, as the actual fourth-order curve varies radically from what would be considered a reasonable extrapolation.

Figures 7 and 8 present the X-band system noise temperatures both with and without atmospheric contribution. Note again two sets of data. The upper data set was taken with the TWM-1 maser, while the lower data set was taken with the TWM-2 maser. Only the upper data set was curve-fitted, owing to the limited elevation angle range of the lower set. It appears that a constant difference (approximately 3 kelvins) separates the two sets. For the X-band curve fit in Fig. 7, the extrapolation is constant above 75 degrees; for Fig. 8 it is constant above 85 degrees.

IV. Error Analysis

A comprehensive review of error sources in this antenna calibration scheme is given in [1] and [2]. It bears repeating that the major contributor to the error in determination of antenna efficiency is the uncertainty regarding radio source flux density. It is estimated that this uncertainty at S-band is ± 0.3 dB (3σ) (± 3 percent, 1σ); at X-band it is ± 0.5 dB (3σ) (± 4 percent, 1σ). Indeed, the 3 percent efficiency adjustments at S- and X-band are perhaps indicative of this problem. The absolute accuracy in the determination of antenna efficiency for the DSS 14 antenna (as determined similarly for DSS 43 and DSS 63) is thus stated as:

S-band: ± 0.4 dB (3σ)

X-band: ± 0.6 dB (3σ)

V. Comparison of Measured and Expected Antenna Efficiencies

As described in [1] and [2], a comparison was made among the 64-m X-band antenna performance expectations as given by the physical optics (PO) and geometrical theory of diffraction (GTD) programs. The PO analysis was described in great detail in those references and will not be repeated here. Of interest here are the GTD calculations of antenna efficiency as a function of elevation angle, taking into account the long-period (~ 1 to 30 meters) gravitational deformation of the main reflector surface. The GTD-generated efficiency curve is

modified by known or postulated hardware loss (0.821 dB; cf. Table 3, items 8-11 in [1]), and this curve is then further modified by various amounts of so-called Ruze loss, the loss of antenna efficiency due to small-scale (~ 1 - to 100-centimeter) surface roughness.

Figure 9 shows the GTD curves with three levels of surface roughness: 1.06 mm (the design expectation of rms panel and subreflector tolerance), 1.5 mm, and 1.6 mm. It is seen that the DSS 14 antenna efficiency curve corresponds over the entire elevation angle range to an rms surface tolerance of 1.55 mm. Also shown on this curve are the DSS 43 and DSS 63 efficiencies, which, as stated in previous articles, correspond to surface tolerances of 1.7 to 1.8 mm. Note that the DSS 63 antenna differed structurally from the DSS 14 and DSS 43 antennas, and thus the greater efficiency falloff of that antenna is not surprising.

VI. Future Updates for DSS 14, 43, and 63 64-Meter Efficiencies

Due to a recent change in the source size correction factor for radio source 3C274 (see footnote 1), there exists a very small discrepancy in the efficiency values determined for DSS 14 as compared with those determined for DSS 43 and DSS 63. For example, it is possible that the 0.369 percent increase in X-band source size correction might increase the efficiencies of the overseas antennas by as much as 0.2 percent (e.g., from 50.0 percent to 50.2 percent). This increase is small compared to the scatter of data points (cf. Figs. 3 and 4 and Table 1), and thus may be judged as not significant in a statistical sense. Consultation with one of the authors of JPL Report D-3801 (see footnote 1) revealed that the present source size correction value should be considered interim only, and that an updated and highly improved document will be published within the next several months.² His advice in the matter of efficiency adjustment was to make no changes at this time. This article and its two precursors used radio source flux and size corrections which were the latest available at the time. Updated values of efficiency for the three stations can be computed when the future DSN radio source list becomes available, although values of tenths of an efficiency percent are not judged significant in view of the absolute accuracy levels presently attainable.

²M. J. Klein, private communication, Jet Propulsion Laboratory, Pasadena, California.

Acknowledgments

The author wishes to thank the engineers at the Goldstone DSCC, in particular Bill Wood and Tim Gregor, for their efforts in obtaining the data needed for this article. The M&I Group members Jim McCoy and Jerry Crook assisted with much-needed equipment maintenance.

References

- [1] S. D. Slobin, "DSS-63 64-Meter S- and X-Band Efficiency and System Noise Temperature Calibrations," *TDA Progress Report 42-90*, vol. April-June 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 103-115, August 15, 1987.
- [2] S. D. Slobin, "DSS-43 64-Meter S- and X-Band Efficiency and System Noise Temperature Calibrations," *TDA Progress Report 42-90*, vol. April-June 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 116-126, August 15, 1987.

Table 1. Coefficients of second order curve fits for antenna area efficiencies

$$\text{efficiency} = a_0 + a_1\theta + a_2\theta^2$$

where θ = elevation angle, degrees

Coefficient/parameter	S-band (2295 MHz)	X-band (8420 MHz)
	With atmosphere (cf. Fig. 1)	With atmosphere (cf. Fig. 3)
a_0	0.567571	0.434121
a_1	7.446475E-04	2.234165E-03
a_2	-5.887066E-06	-2.082133E-05
Peak efficiency, %	0.59112	0.49405
Peak angle, deg	63.244	53.651
Standard deviation, %	0.00354	0.00853
	Without atmosphere (cf. Fig. 2)	Without atmosphere (cf. Fig. 4)
a_0	0.592390	0.457946
a_1	1.936510E-06	1.572380E-03
a_2	4.675295E-07	-1.541394E-05
Peak efficiency, %	0.594 (see text)	0.49805
Peak angle, deg	(see text)	51.005
Standard deviation, %	(see text)	0.00859

Table 2. Coefficients of fourth-order curve fits for S- and X-band system noise temperatures

$$T_{\text{system}} = a_0 + a_1\theta + a_2\theta^2 + a_3\theta^3 + a_4\theta^4$$

where θ = elevation angle, degrees

Coefficient/parameter	S-band (2295 MHz) ^a	X-band (8420 MHz) ^b
	With atmosphere (cf. Fig. 5)	With atmosphere (cf. Fig. 7)
a_0	4.919551E+01	5.836215E+01
a_1	-2.040157E+00	-2.234962E+00
a_2	6.445888E-02	6.128831E-02
a_3	-9.468854E-04	-7.517988E-04
a_4	5.169310E-06	3.378224E-06
	Note: If $\theta \geq 65$ deg, $T = 21.161$ K	Note: If $\theta \geq 75$ deg, $T = 25.211$ K
Zenith noise temperature, K	21.161	25.211
Standard deviation, K	0.334	0.358
	Without atmosphere (cf. Fig. 6)	Without atmosphere (cf. Fig. 8)
a_0	2.727111E+01	3.355993E+01
a_1	-2.675004E-01	-5.586415E-01
a_2	2.027795E-03	1.278769E-02
a_3	2.917834E-05	-1.356839E-04
a_4	-3.911682E-07	5.376014E-07
	Note: If $\theta \geq 70$ deg, $T = 19.098$ K	Note: If $\theta \geq 85$ deg, $T = 23.203$ K
Zenith noise temperature, K	19.098	23.203
Standard deviation, K	0.239	0.284

^aS-band (2295 MHz) specs: Maser SPD Blk IV, S/N 4002, 1.83 K; Maser Blk V, S/N 5002, 4.48 K.

^bX-band (8420 MHz) specs: Maser Blk II, TWM-2, S/N 2007, 3.46 K; Maser Blk IIA, TWM-1, S/N 2011, 3.82 K.

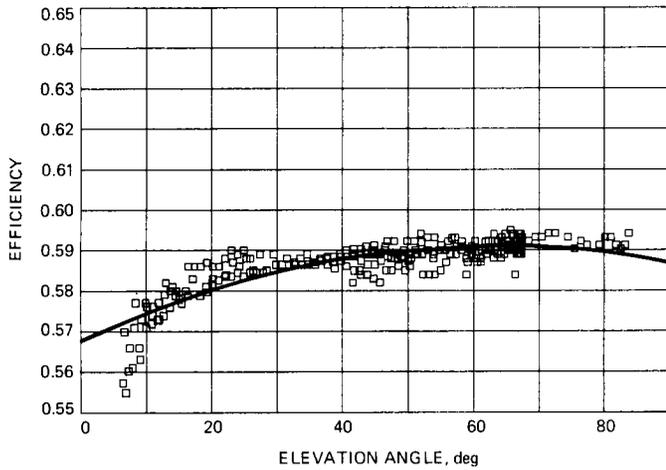


Fig. 1. DSS 14 64-m S-band (2295-MHz) area efficiency with atmospheric attenuation included

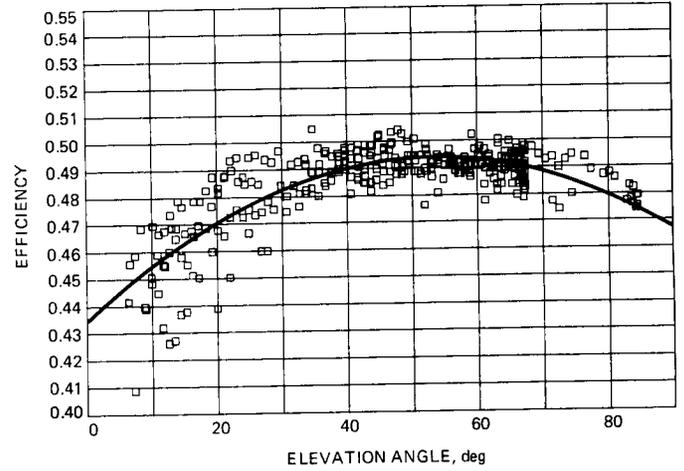


Fig. 3. DSS 14 64-m X-band (8420-MHz) area efficiency with atmospheric attenuation included

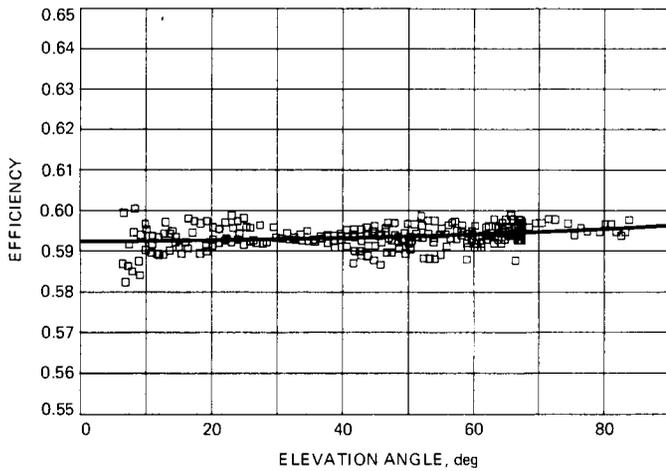


Fig. 2. DSS 14 64-m S-band (2295-MHz) area efficiency without atmospheric attenuation

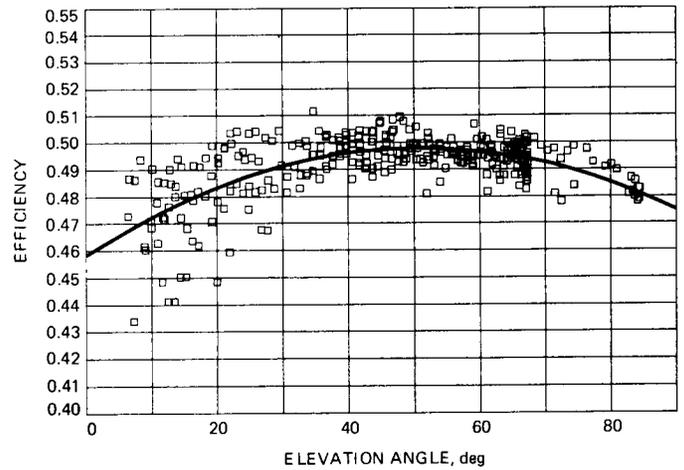


Fig. 4. DSS 14 64-m X-band (8420-MHz) area efficiency without atmospheric attenuation

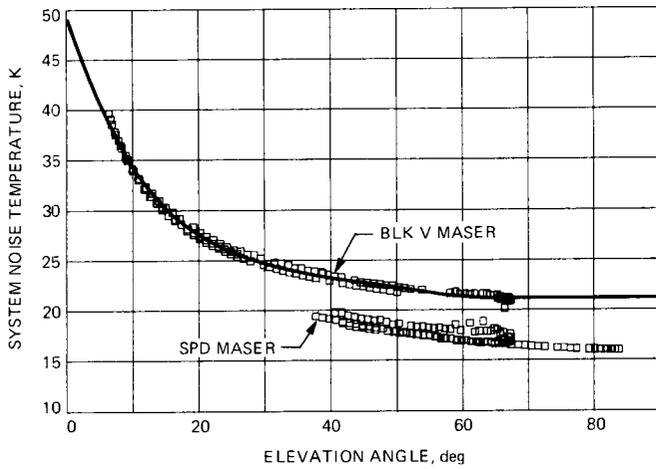


Fig. 5. DSS 14 64-m S-band (2295-MHz) system noise temperature, including atmospheric contribution

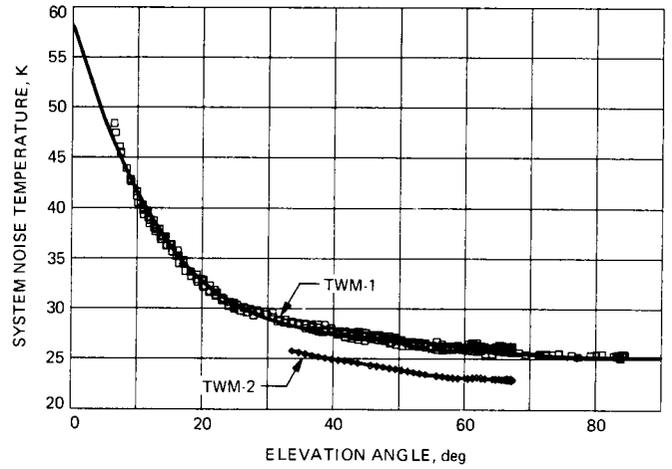


Fig. 7. DSS 14 64-m X-band (8420-MHz) system noise temperature, including atmospheric contribution

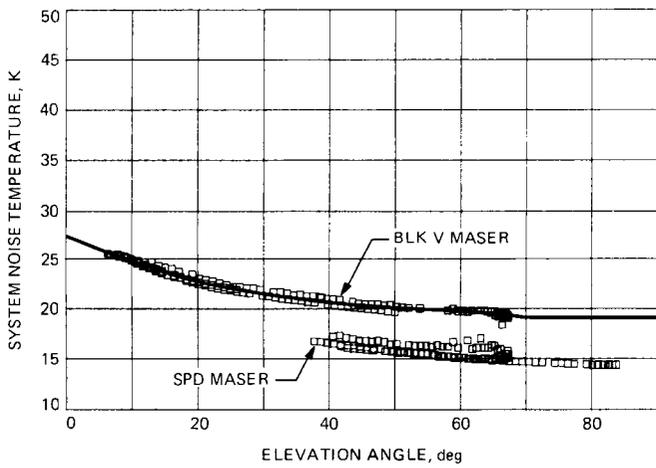


Fig. 6. DSS 14 64-m S-band (2295-MHz) system noise temperature without atmospheric contribution

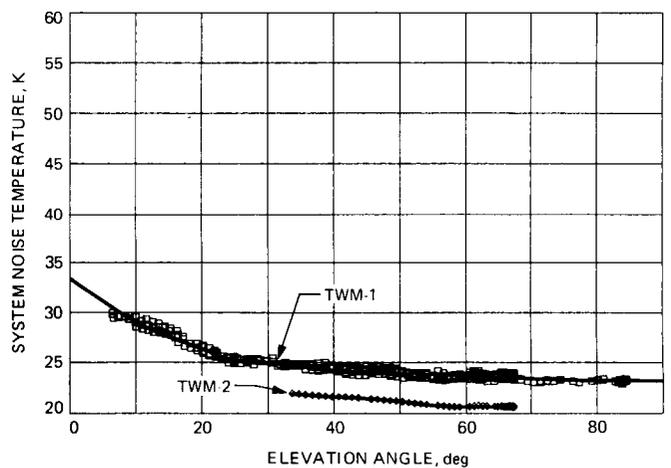


Fig. 8. DSS 14 64-m X-band (8420-MHz) system noise temperature without atmospheric contribution

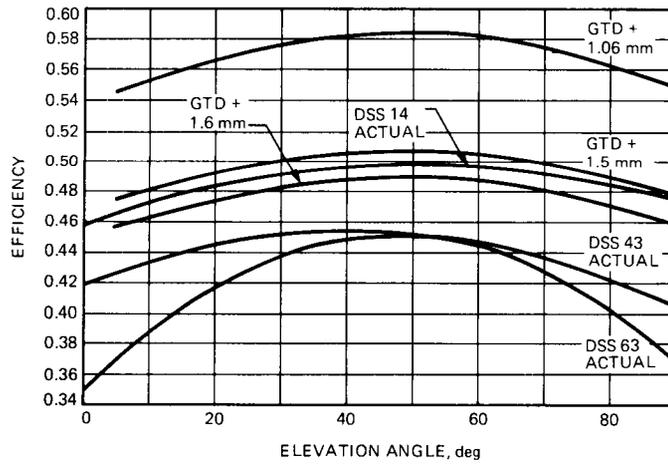


Fig. 9. Comparison of DSS 14, 43, and 63 64-m X-band measured efficiency and GTD-calculated efficiency for various Ruze surface tolerance values, without atmosphere

New Mode Switching Algorithm for the JPL 70-Meter Antenna Servo Controller

J. A. Nickerson

Ground Antenna and Facilities Engineering Section

This article describes the design of new control mode switching algorithms and logic for JPL's 70-m antenna servo controller. The old control mode switching logic was reviewed and perturbation problems were identified. Design approaches for mode switching are presented and the final design is described. Simulations used to compare old and new mode switching algorithms and logic show that the new mode switching techniques will significantly reduce perturbation problems.

I. Introduction

The servo controller for the NASA/JPL 70-m-diameter antenna contains three control algorithms by which antenna position is controlled. The control algorithms provide control for antenna slewing, computer tracking, and precision tracking.

Switching between control algorithms plays a key role in tracks which require frequent slewing between target positions. Mission support and particularly Very Long Baseline Interferometry (VLBI) tracks require frequent slewing, for which antenna repositioning time is critical. Thus, not only must the control algorithms be designed to minimize antenna repositioning time, but switching between the control algorithms must be optimized to minimize perturbations.

The transition between control algorithms must be smooth to prevent acceleration/rate perturbations which can excite structural resonances, thereby increasing repositioning time and adding to mechanical wear. Reducing perturbations is particularly important for the 70-m antenna because of its low structural resonances and existing mechanical gear wear.

This article describes new mode switching logic and algorithms which will minimize acceleration/rate perturbations in previous servo controllers.

II. Background

The antenna servo control system consists of a position loop closed around a rate loop. The rate loop is an analog type I controller. A rate command (voltage) from the position loop controller is compared to filtered tachometer feedback rate (voltage) to create a rate error. The rate loop adjusts the actuator command signal to null the rate error which results in moving the antenna at the commanded rate.

The position loop is closed by a digital computer, the Antenna Servo Controller (ASC). The ASC positions the antenna based on predicts (position commands). Three control algorithms reside in the ASC: (1) a slew mode called the Large Error mode; (2) a computer tracking mode called the Small Error mode; and (3) a precision tracking mode called the Precision mode. The Large Error mode is used to slew

the antenna over large angular displacements. The Small Error mode is used to track predicts with encoder feedback. The Precision track mode is used to track a precision positioner, the master equatorial (ME), using an optical link (an autocollimator) as a position feedback device.

In the old Large Error mode, a digital rate servo was used to slew the antenna. The input to the rate servo was calculated from predicts and accounted for target motion. The rate servo was a state controller where the state variables were estimated by a constant gain state estimator. Encoder feedback was used to update the state estimator.

The old Small Error mode was a type II position controller. The Small Error mode also positioned the antenna based on predicts. A state controller with an added integral error state was used for the type II position controller. The state variables were estimated by a constant gain state estimator. On-axis encoder feedback was used to update the state estimator and ultimately position the antenna.

The old Precision mode used a type II position controller similar to the Small Error mode. Position feedback was provided by an autocollimator instead of on-axis encoders. The autocollimator measures optical misalignment between the antenna and the master equatorial and produces a voltage proportional to the position error. The autocollimator signal, once sampled, was digitally filtered and a secant correction was applied to the azimuth position error for high elevation angles. The autocollimator error signal was used to calculate the integral error and the position error states in the state estimator. The rest of the estimated states were calculated from the antenna's on-axis encoder feedback. Using encoder feedback for estimating the other state variables improved damping.

Each of the above control algorithms calculated a rate command needed to drive the rate loop. The rate command was limited before being converted to an analog voltage. Both rate and acceleration limiters were used to limit the rate to ± 0.25 degree/sec and the acceleration to ± 0.20 degree/sec².

Switching between the three control modes was based on the magnitude of rate commands, position errors, and autocollimator acquisition. Figure 1 presents the old switching logic state diagram. The initial state was the Small Error algorithm. A transition to the Large Error algorithm occurred when the absolute value of the new rate command was greater than a predetermined upper limit ($|\text{NEW_RATE}| > \text{U_LIMIT}$). The transition from the large to the Small Error algorithm occurred when the position error was less than the error limit ($|\text{POS_ERROR}| < \text{LARGE_ERROR}$).

Switching between the Small Error and Precision modes was more complicated. Three conditions were needed prior to switching to the Precision mode from the Small Error mode: (1) the Precision mode was commanded (i.e., $\text{PRECISION_MODE} = \text{'TRUE'}$); (2) the autocollimator was within signal acquisition range (i.e., $\text{ACQUISITION} = \text{'TRUE'}$); and (3) the digitally filtered autocollimator signal was small (i.e., $|\text{F_ERROR}| < \text{FILTER_LIMIT}$). All three conditions had to be met to switch to the Precision mode from the Small Error mode.

Conversely, when one of the following three conditions was satisfied, a switch was made from the Precision mode to the Small Error mode: (1) the Small Error mode was commanded ($\text{PRECISION_MODE} = \text{'FALSE'}$); (2) autocollimator acquisition was lost ($\text{ACQUISITION} = \text{'FALSE'}$); or (3) the absolute value of the calculated rate command was greater than an upper limit (i.e., $|\text{NEW_RATE}| > \text{U_LIMIT}$). Switching from the Large Error mode to the Precision mode and vice versa was an illegal state transition.

III. Undesirable Perturbations

The old switching logic caused rate perturbations when switching between control algorithms. Figure 2 is a strip chart record of the rate command for the old switching sequence from Small to Large and from Large to Small Error modes. The data represent the rate command measured at the output of the D/A (digital to analog) converter. Initially, the Small Error algorithm was holding position. A one degree position command created the condition ($\text{NEW_RATE} > \text{U_LIMIT}$) causing a switch into the Large Error mode. The rate command increased linearly to the maximum rate because of acceleration limiter saturation. As the antenna slewed, the position error was reduced. When the position error became small enough a switch was made to the Small Error mode. A transient rate command occurred because of gain differences between Small and Large Error modes.

The old switching sequence from Small to Large and from Large to Small Error modes caused undesirable perturbations. The rate perturbations are excessive and can excite structural resonances and increase gear wear.

The old switching sequence from Small to Precision and from Precision to Small Error modes was simulated. Figure 3 presents the results of the transition based on a 30 millidegree misalignment between the antenna and the ME. The rate command measured at the D/A converter was plotted versus time. Initially, the Small Error mode held position. A switch was made into Precision mode. After five seconds, a second switch was made back to Small Error mode. Simulations showed that instability existed due to misalignment errors and differences

in the gain vectors between Small and Precision modes. Misalignments greater than 20 millidegrees cause instability and antenna oscillations. Misalignments less than 50 millidegrees would oscillate but eventually become stable.

The old switching sequence from Small to Precision and from Precision to Small modes caused undesirable perturbations. The transitions exhibit unstable response for large misalignment errors and perturbations which can excite structural resonances and increase gear wear.

The design goal was to develop new switching algorithms and logic which minimize slew time and rate perturbations and provide stable transitions. Therefore, the settling time required to move the antenna between targets through large angle differences will be minimized for VLBI tracking. Also, perturbations will be minimized to reduce structural resonance excitation and drive gear wear.

IV. New Design Approach

The design approach was to redesign the transition algorithms and logic. Switching between Small and Large Error modes and that between Small and Precision Error modes were treated separately. Note that the basic tracking control strategies, acceleration limiting, and rate limiting were not changed.

A. Switching Between Small and Large Error Modes

Analysis of switching logic between Small and Large Error modes showed that transition perturbations could best be improved by developing a new Large Error mode. Several control strategies for replacement of the Large Error mode were investigated, i.e., rate servos, proportional controllers, gain scheduling, and nonlinear gains. The control strategies were analyzed for smooth mode transitions, response time, stability, and simplicity. Design iterations suggested that a two-part Large Error algorithm provided the "best" switching strategy in terms of stability. First, a new Large Error mode will accelerate the antenna from the Small Error mode to the maximum slew rate. A new Modified Small Error algorithm, which will be part of the Small Error mode, will decelerate the antenna from the maximum slew rate to Small Error algorithm tracking.

Accelerating from the Small Error mode to maximum slew rate will be accomplished by using a predetermined acceleration profile. This profile will minimize rate perturbations, provide known acceleration time, and smooth transitions.

Decelerating from maximum slew rate to Small Error mode tracking will be accomplished by a modified Small Error algorithm which eliminates the integral error state. The Modified Small Error algorithm would then be a type I controller. Transition to the type II Small Error algorithm is accomplished by adding the integral error state. The transition to the Small Error mode from the modified algorithm will be smooth and stable with minimal settling time.

B. Switching Between the Small Error and Precision Modes

Both gain matching between Small and Precision modes and new switching algorithms were investigated to minimize rate perturbations during mode switching. Matching small and precision mode state feedback gain vectors improved transition performance but did not improve transition stability. Thus, a switching algorithm was needed. Potential transition algorithms included gain scheduling, a rate servo, and filtering the position error. Investigations showed that filtering position error significantly improved stability and provided smooth mode switching.

V. Detailed Design of Mode Switching Algorithms

The new mode switching algorithms and logic were designed to switch between Small and Large Error modes and between Small Error and Precision modes. The switching logic determines when to switch to a different control algorithm while the mode switching algorithms control the transition between control algorithms.

The new 70-m antenna control algorithm switching logic is presented in Fig. 4. The Small Error mode is the initial state in the state diagram. Switching to the Large Error mode from the Small Error mode occurs when a new position command requires the antenna to move over a large angle. New position commands are received and evaluated once per second to determine if a slew is required. Transitions from the Large to the Modified Small Error algorithm occur when the maximum slew rate is attained. All transitions between the Small Error and Precision modes remain the same as the old transition logic.

Three new mode switching algorithms were developed: a Large Error algorithm, a Modified Small Error algorithm, and a digital filter between Small Error and Precision mode transitions. The new Large Error algorithm has two parts: (1) to accelerate the antenna to the maximum slew rate (this is called the new Large Error mode); and (2) to decelerate the antenna from maximum slew rate and transition into the Small Error algorithm (this is part of the new Small Error mode).

The purpose of designing a new Large Error mode is to smoothly accelerate the antenna to the maximum slew rate. When a slew is required, a predetermined acceleration profile is used to accelerate the antenna from the tracking rate to the maximum slew rate in the appropriate direction.

The acceleration profile is generated in real time by using a digital filter. Initializing the filter with the last rate command and commanding the maximum rate produce a step response referenced to filter initialization. A third order Bessel digital filter was used as the Large Error algorithm rate command filter. The advantages are twofold: a small overshoot due to a step in rate and a smooth acceleration profile. The bandwidth of the filter is selected such that the ratio of maximum rate to maximum acceleration is less than or equal to the ratio of the rate limit to the acceleration limit. The 70-m antenna has a rate limit of ± 0.25 degree/sec and an acceleration limit of ± 0.20 degree/sec². Thus, the maximum acceleration to rate ratio must be less than or equal to 0.8.

Figures 5 and 6 show the rate step response and its derivative, respectively, for a Bessel filter with a 1 rad/sec bandwidth. The step response shows how the rate command (filter output) accelerates to the maximum rate and has a 4 percent overshoot. The derivative of the step response (in Fig. 6) shows the acceleration profile by which the antenna is accelerated. The acceleration to rate ratio is 0.8.

Once the maximum rate is achieved in the Large Error mode, a transition is made to a type I controller. For simplicity, the tracking Small Error algorithm is used for the type I controller with the integral error limited to zero which effectively "turns off" the integrator. As a result, the tracking controller acts as a type I controller and smoothly decelerates the antenna. When the position error is small (a 30 millidegree threshold is used), the integral error is "turned on" and tracking with a type II controller begins. Setting the integral error limit to a finite value prevents oscillations from occurring when tracking rates vary.

The transitions between Small Error and Precision modes were also improved by digital filtering. The same Bessel filter described above was used to filter the position and autocollimator errors during transitions. Figure 7 describes the Small-to-Precision and Precision-to-Small-Error mode state switching diagram. When a Small Error to Precision transition is made, the mode switching filter is initialized with the last position error. The new position error (or equivalently the autocollimator error) is calculated. The new position error is filtered by the transition filter. This filtered error is used by the precision control algorithm to position the antenna. When the filtered error is greater than the position error, the filter is "turned

off." The same procedure is used when switching from the Precision mode to the Small Error mode.

VI. Simulation Results

The new transition algorithms and logic were coded in the 70-m antenna simulation software and simulated using a Monte Carlo simulation. Mode switching simulations between Small and Large Error modes and between Small and Precision modes were made.

A mode switching simulation was made between the new Small and Large Error modes to demonstrate switching logic and algorithms. Figure 8 presents the simulated rate command measured at the output of the D/A converter. Initially, the Small Error algorithm was holding position. A one degree position command caused a switch into the Large Error algorithm (the new Large Error mode). The rate command smoothly accelerated to the maximum rate. At the maximum rate, the transition was made to the Modified Small Error algorithm (part of the new Small Error mode). No rate perturbations occurred at the transition because the rate command produced by the Modified Small Error algorithm is greater than the maximum slew rate and is thus limited to the maximum slew rate. The rate limiter effectively eliminates transition perturbations.

As the position error decreased, the rate command dropped below the maximum slew rate. After the position error became small enough, a transition was made to the Small Error algorithm by "turning on" the integral error state.

Comparing Fig. 8 to the results of the old mode switching logic in Fig. 2 shows that a significant improvement was attained in the number of switching perturbations. The only discontinuity produced by the new mode switching occurs when initial deceleration from the maximum slew begins. At this point, the acceleration is discontinuous and has a jump which, at most, is equal to the acceleration limit. However, this discontinuity is acceptable.

The new mode switching algorithm between Small Error and Precision modes was also simulated. Figure 9 presents the simulated rate command measured at the output of the D/A converter. Initially, the Small Error algorithm was holding position. A transition was immediately made to the Precision mode. A second transition was later made from Precision to the Small Error mode. A 30 millidegree misalignment error between the antenna and the ME was assumed for both transitions. Simulations showed that transition instability due to misalignment errors was reduced. Misalignment errors of up to 50 millidegrees were simulated and were stable. With 50

millidegree misalignments the acceleration limiter started to become saturated but the control algorithm remained stable.

VII. Summary

New mode switching algorithms and logic were designed for switching between 70-m control algorithms. Transitions to the Large Error algorithm are made when the position command requires a large angular motion. A transition from the Large to Small Error mode is made when the position error is small. All transition logic between the Small Error and Precision modes remains the same as the old mode switching logic.

A new Large Error algorithm was designed. The Large Error algorithm was separated into two parts: the Large Error mode and the Modified Small Error algorithm. The Large Error mode accelerates the antenna from the Small Error mode to the maximum slew rate by using a digital Bessel filter. The Modified Small Error algorithm (which is also part of the new

Small Error mode) is a type I controller. It decelerates the antenna from the maximum slew rate and transitions into the type II Small Error algorithm by "turning on" the integral error state in the state controller. Simulations showed the new mode switching between Small and Large Error modes provide smoother transitions with fewer acceleration/rate perturbations.

The mode switching between Small Error and Precision modes was improved by using the same Bessel filter as in the Large Error mode and by matching control gain. The misalignment error was filtered to provide smoother and more stable transitions. Simulations of the Small and Precision mode switching showed improved transition stability, robustness, and smoothness.

The new servo controller mode switching logic and algorithms for the 70-m antenna provide better overall performance than the old mode switching logic. The better performance will reduce potential structural resonance excitation and gear wear.

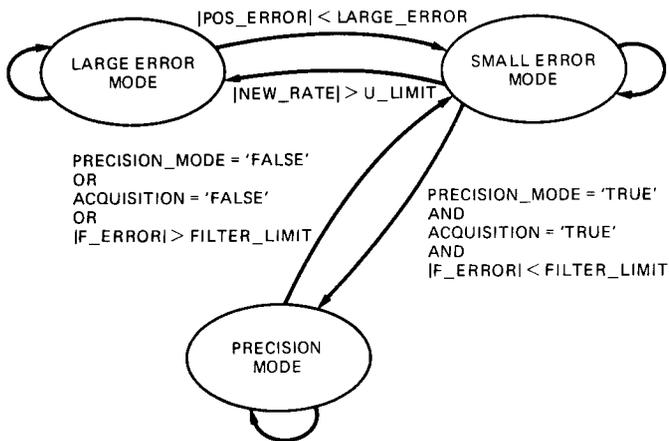


Fig. 1. Old control mode switching logic

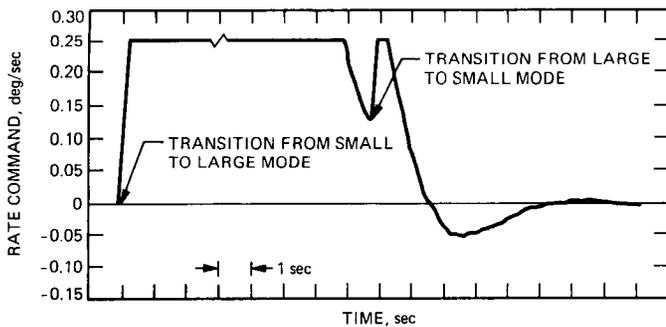


Fig. 2. Mode switching between Small and Large Error modes

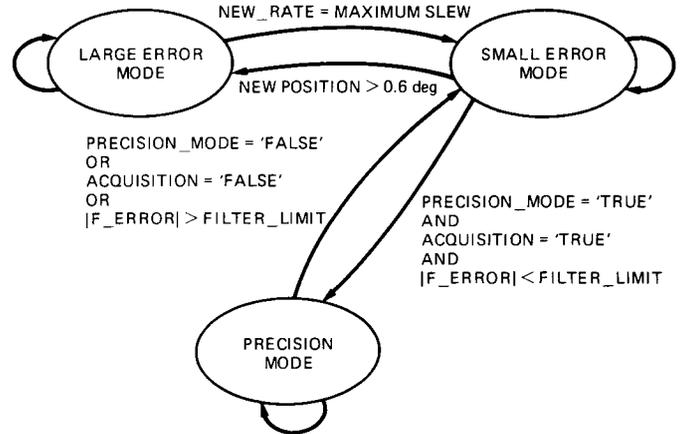


Fig. 4. New 70-m control mode switching logic

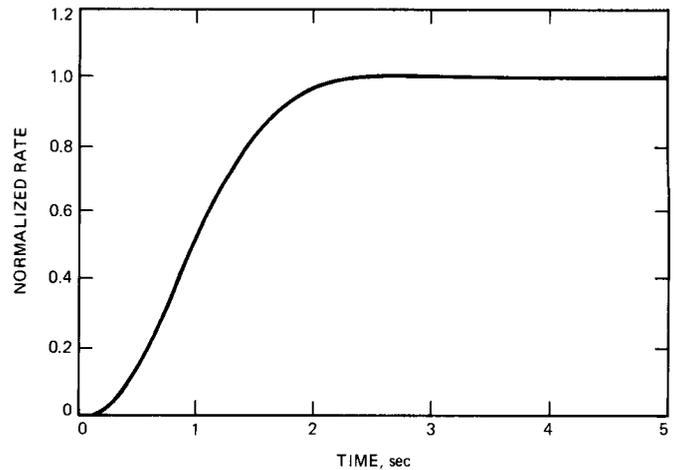


Fig. 5. Normalized Bessel filter step response (third-order Bessel filter; 1 rad/sec cutoff frequency)

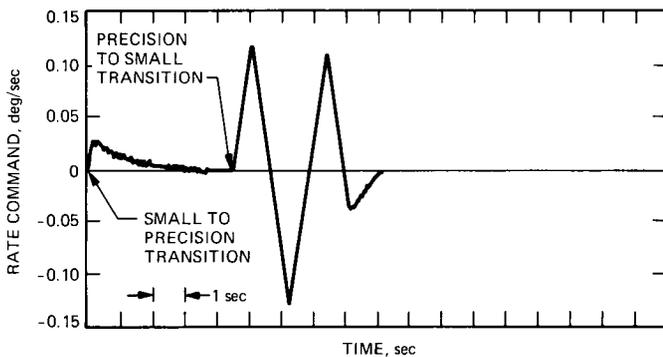


Fig. 3. Mode switching between Small Error and Precision modes (30-millidegree misalignment assumed)

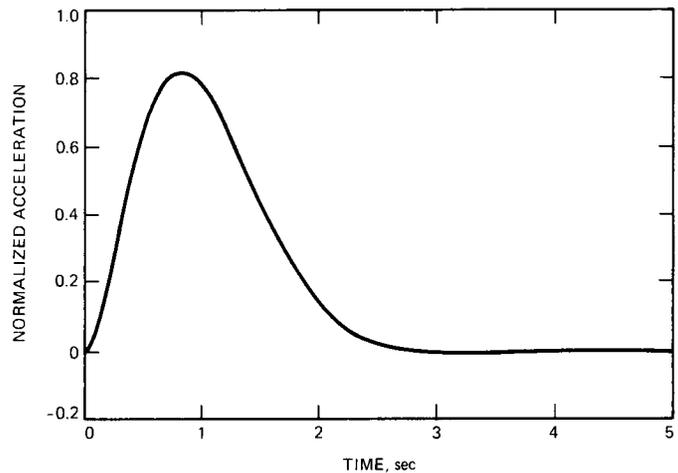


Fig. 6. Derivative of normalized Bessel filter step response (maximum acceleration is 80 percent maximum rate)

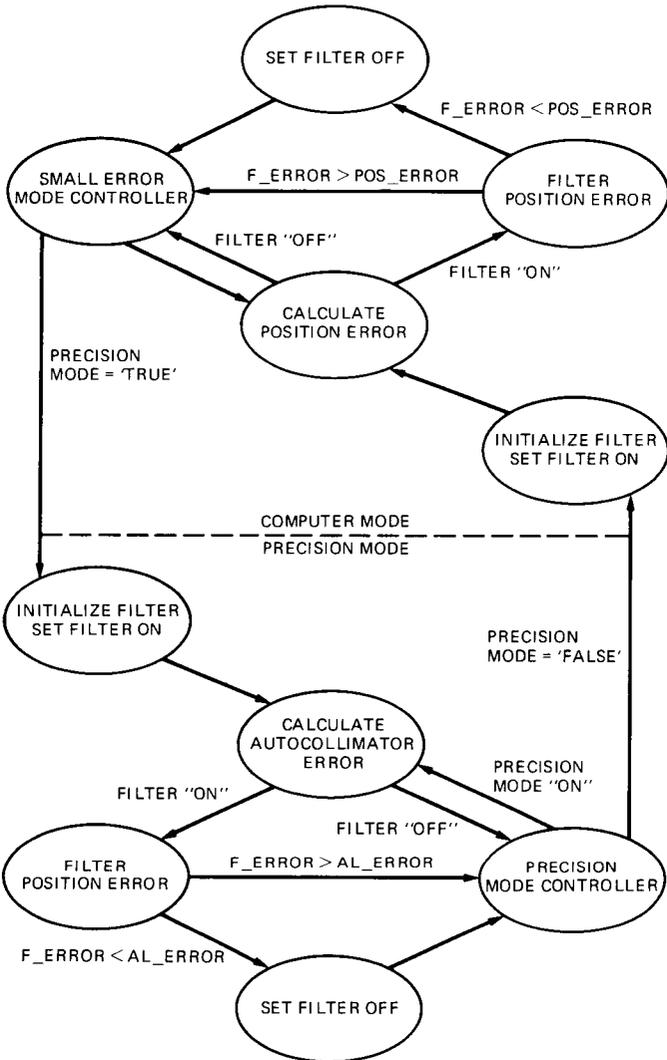


Fig. 7. New 70-m mode switching logic between Small Error and Precision modes

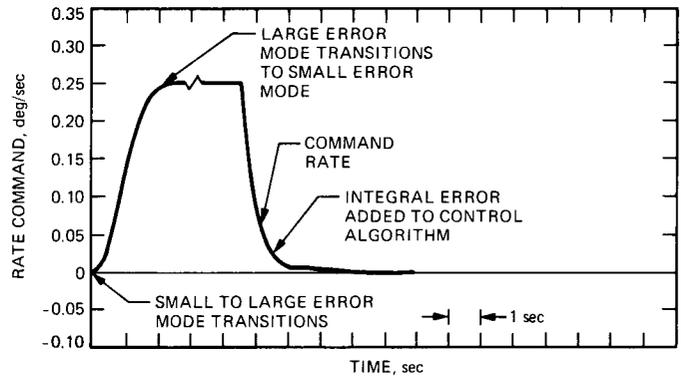


Fig. 8. New 70-m mode switching between Small and Large Error modes

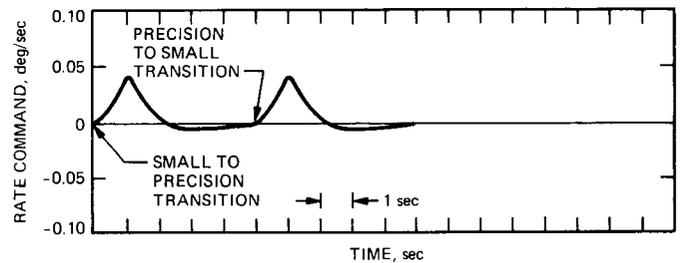


Fig. 9. New 70-m mode switching between Small Error and Precision modes (30-millidegree misalignment assumed)

518-74
129234

N88-18792

120

Optical Links in the Angle-Data Assembly of the 70-Meter Antennas

M. D. Nelson, J. R. Schroeder, and E. F. Tubbs
Automated Systems Section

In the precision-pointing mode the 70-meter antennas utilize an optical link provided by an autocollimator. In an effort to improve reliability and performance, commercial instruments have been evaluated as replacement candidates, and upgraded versions of the existing instruments have been designed and tested. The latter have been selected for the Neptune encounter, but commercial instruments with digital output show promise of significant performance improvement for the post-encounter period.

I. Introduction

When the 70-meter antennas of the Deep Space Network are operated in the precision-pointing mode, an optical link provided by an autocollimator plays an important role in the measurement of pointing angles. Over the past several years there has been an effort to upgrade both the reliability and the performance of the precision-pointing mode by upgrading or replacing the instrumentation which provides this link. Section II reviews the rationale for the use of a pointing subsystem requiring an optical link and describes the elements of the subsystem. Section III discusses the general technology of autocollimators and the parameters used to characterize them. It then describes and compares detection techniques used in commercial autocollimators. Section IV reports on the evaluation of several commercial instruments. Most of these evaluations are the result of testing in the JPL autocollimator facility. Section V describes development work done at JPL in the areas of improvements to the present autocollimators and the interfacing of an autocollimator with a digital output to the

antenna control system. Section VI summarizes the present situation and makes recommendations concerning the autocollimators both for the near term and for the period after Neptune encounter.

II. Rationale for the Optical Link in Antenna Pointing

The 70-meter antennas, like most large antennas, are in altitude-over-azimuth mounts. In the ideal situation an alidade structure rotates about an axis perfectly aligned with the local vertical and supports a perfectly horizontal axis about which the tipping structure rotates carrying the dish, subreflector, feed, etc. In this ideal situation it is only necessary to translate the astronomical coordinates, declination and hour angle, into azimuth and elevation in order to point the antenna. In actual practice the situation is far from ideal. Foundations may settle, moving the azimuthal axis off the vertical and

giving a tilt to the elevation axis which depends upon the azimuthal angle. In addition, solar heating, aerodynamic forces, and gravitational loading will distort the structure, introducing pointing errors.

There are two general approaches to dealing with the situation. One is to measure rotation about the azimuth and elevation axes with shaft encoders and depend upon mathematical modeling both for the correction of systematic errors and disturbances and for the transformation to declination and hour angle. The other approach is to utilize a separate pointing-metrology system which is designed to be less susceptible to errors and disturbances than the antenna and which determines the pointing direction of the antenna. Such a pointing-metrology system can range in complexity from one which surveys many points on the dish, subreflector, and feed and solves for the instantaneous pointing direction to one which determines a single reference direction on the tipping structure. Neither approach can completely eliminate the need for modeling, but they can reduce the number of parameters which must be modeled. The present system for the 70-meter antennas determines a single direction, that of the normal to the intermediate reference structure located on the center line of the dish between the vertex of the dish and the elevation axis.

There are five major elements in a pointing-metrology system of the latter type: a mechanical support structure that is as independent of the rest of the antenna structure as possible; an element articulated about two axes which establishes the pointing direction; encoders to determine the pointing direction of the articulated element; a link to the reference element on the antenna; and an enclosure to minimize environmental disturbances. In principle the link could be provided in a number of ways. In practice there are strong reasons for making it an optical one. The chief reason is disturbance isolation. By its nature an optical link is unidirectional and, in the appropriate configuration, isolates the articulated element from disturbances to the antenna, particularly vibration. By using an optical link as opposed to a mechanical one it is possible to have independent motion of the antenna and the articulated element. This facilitates slewing and allows for operation in a lower-precision mode with elevation and azimuth encoders, should the precision mode be inoperative. A light beam is weightless and can, with proper design of the instrumentation, transfer a pointing direction free from errors caused by gravity deflections. Finally, the techniques for providing the instrumentation are well established with substantial field experience in a variety of applications.

A metrology system, such as described above, may be configured to operate in two distinctly different ways. The artic-

ulated element of the metrology system may be pointed in the desired direction with its encoders and the link used to generate a control signal for the antenna, causing it to follow the articulated element. Alternatively, the link may be used to drive the articulated element to follow the antenna. The encoders of the metrology system then give the present point direction of the antenna. The pointing-metrology system of the 70-meter antennas is used in the first of these two modes although much of the hardware needed for the second mode is in place.

The actual configuration for the pointing metrology system of the 70-meter antennas, shown schematically in Fig. 1, is the following: A central support column rises from the foundation along the azimuthal axis to a point a few feet below the elevation axis. The articulated structure is in an equatorial-fork mount similar to those used for telescopes and supports a plane mirror. The pointing direction of the metrology system is the normal to the plane mirror. Its angular position in declination and hour angle is measured by an Inductosyn transducer on each axis. The pointing direction is transferred to the intermediate reference structure by an optical link provided by an autocollimator. The autocollimator projects a beam of light onto the plane mirror and develops an electrical output signal proportional to the angular offset of the returned beam. These signals are developed for rotating about the elevation axis and a direction orthogonal to it (cross elevation) and serve as the control signals for pointing in elevation and azimuth. The signal flow is shown in Fig. 2. The critical nature of the optical link is evident and the following sections discuss its implementation in detail.

III. Autocollimator Technology

A. General Characteristics

A general autocollimator configuration is shown in Fig. 3. Light emitted from a source at the focus of a lens is reflected by a beam splitter and formed into a parallel beam by the lens. This beam is returned by the mirror whose rotation is being measured and is focused by the lens on a position-detection element. When the mirror is precisely perpendicular to the axis of the lens, the returned spot is centered on the element. As the mirror is rotated the spot is moved on the element. This generates a signal proportional to the linear displacement of the spot which for the small angles involved is proportional to the mirror rotation. The angular coverage at zero working distance is determined by the ratio of the field-stop size to the focal length of the objective lens. As the working distance increases, a given angular deflection of the mirror causes more of the return beam to miss the entrance aperture of the lens. This results in a decrease in signal which limits performance.

B. Parameters

Several parameters are required to characterize an autocollimator. Some of these can be represented on the transfer curve shown schematically in Fig. 4, which plots the electrical output as a function of angular offset. Of particular importance for control functions are those parameters relating to the central part of the curve: the slope, the linearity, and the offset at null. The other parameters, the linear range and the acquisition range, are of importance for initial alignment but do not figure in the performance once a control loop has been closed. Other parameters of importance not shown on the curve are maximum working distance, noise-equivalent angle, electrical phase shift, cross-coupling, and all-attitude sensitivity. Table 1 gives the specified values for these parameters for the DSN autocollimators.

The maximum working distance and the noise-equivalent angle (NEA) are very closely connected. At the working distances encountered in the antenna installation the NEA is determined primarily by air turbulence and not by the intrinsic instrument characteristics. At very short working distances the converse is true and the NEA is determined by the instrument characteristics. The electrical phase shift is a measure of the response to a moving mirror. It must be traded off against instrumental NEA. The cross-coupling is a measure of the degree to which an angular offset about one mirror axis produces a spurious output for the other axis. All-attitude sensitivity is a measure of spurious outputs produced by tipping the instrument in a gravitational field.

C. Description of Detection Techniques

The selection of the technique for detecting the position of the returned spot in the focal plane is a critical design decision, and a number of techniques have been used in commercial instruments over the years. Figures 5 and 6 show the techniques used in instruments tested in this study. Figure 5 shows mechanical scanning of the image of the source aperture over the detector field of view. There are many ways in which this can be done both in terms of the shapes of the source and detector apertures and in terms of the scan pattern. The particular one shown is used in the Kollmorgen-built instruments currently being used on the antennas. A rotating optical element nutates the image of the source aperture on the detector aperture. In the nutational motion the center of the image travels in a circle, but the image does not rotate. When the returned beam is in the null position this circular motion is centered on the detector aperture and the detector output has zero first-harmonic content. As the image is displaced from the center of the detector, the first harmonic of the detector output increases. The phase of this signal indicates the direction of displacement. Since the signals for each axis are in quadra-

ture they can readily be separated by phase-sensitive rectification. The square source aperture generates a strong fourth harmonic in the detector output, but this is removed by the filtering which follows the phase-sensitive rectification.

Figure 6 shows four detecting techniques which do not involve mechanical scanning. In the interest of clarity all configurations are shown for a single axis of motion. All are adaptable to two-axis operation but not with equal facility. The dual source has been widely used in a variety of implementations for many years. The sources are half-wave modulated 180 degrees out of phase, and synchronous rectification of the ac component of the detector output is used to develop the output signal. Davidson Optronics has used this technique for a number of years in their line of autocollimators. They use gas-discharge lamps as the light source and implement two-axis operation by using two wavelength regions: one obtained from mercury-vapor lamps, and the other provided by neon lamps.

The dual source has also been used by Micro-Radian for several years in their line but in a different implementation [1]. They use two extended-area light emitting diodes (LEDs) and avoid the light losses inherent in a beam splitter by placing the detector behind the gap between the sources. Since this arrangement does not lend itself to two-axis operation with a single optical system, Micro-Radian uses two separate optical systems mounted close together. Cross talk is avoided by using a different modulation frequency for each axis.

The split detector is used in a variety of optical sensing systems including autocollimators and is commonly constructed by placing two silicon photodetectors on a single substrate. The outputs of the two detectors are combined as $(A - B)/(A + B)$ to yield the position-sensitive signal. Generally the dc processing is used. Unlike the dual source, the dual detector is readily extended to two-axis operation by dividing the detector into quadrants and processing the signals as $(A + B - C - D)/(A + B + C + D)$ for one axis and $(A + B + C - D)/(A + B + C + D)$ for the other axis.

Lateral-effect photodetectors are an alternative use of silicon-detector technology. The detectors are produced by diffusing a P-type dopant and an N-type dopant into opposite sides of intrinsic silicon base material. Two contacts, one at each end of the sensor, provide alternative paths for the photocurrent, and the division of the current between the two contacts is a measure of the location of the spot of light on the detector. Linearity is fairly good over the central 25 percent of the area and the total range can be quite large [2]. Like the dual detector the processing is $(A - B/A + B)$, and extension to two axes is accomplished by adding a second set of contacts.

The linear-array CCD (charge coupled device) is significantly different from the other devices of Fig. 6 in that it involves discrete pixels read out sequentially. Typically the source aperture is made in the form of a slit so that motion in the direction perpendicular to the array will not cause loss of signal. The slit image is made wide enough to spread the light over several pixels for accurate centroid determination.

D. Comparison of Detection Techniques

A mechanical scanning system can, if carefully designed and constructed, yield high quality performance. The accuracy of the scanning motion must be high, and the sensitivity of the detector and illumination of the source aperture must be very uniform. Since there are practical upper limits to the angular speed of the scanning element, the modulation frequency is limited to the 100 to 200 Hz range. The filtering required after the synchronous rectification in this frequency range limits the signal bandwidth for some applications. The use of mechanical scanning also imposes an operational limitation because of the finite bearing lifetime.

The dual source technique also requires good uniformity of illumination over the source apertures. In addition the source must have good modulation characteristics. In practice this also means sources of low brightness if a reasonable level of system complexity is to be maintained. The gas discharge lamps used by Davidson Optronics for this purpose have only a few hundred hours of life. The LEDs used by Micro-Radian have very long life but low brightness. This limits their application to single-axis instruments and modest working distances.

The dual or quadrant detector provides a high-accuracy system over a limited angular range. Again, source uniformity is important. If a substantial linear range is required the source must be of large lateral extent as well as very uniform. If the linear range is limited but a large acquisition range is required, this detector is very good. Since the zero position is marked by the physical divisions of the detector surface, the device lends itself to systems requiring a well-defined zero.

In contrast, the lateral-effect photodiode can be used with a concentrated source which need not be particularly uniform since charge is integrated over the entire source image. An additional advantage of the lateral-effect photodiode is that it intrinsically has a large acquisition range. Disadvantages are that the central region over which the behavior may be considered linear is very limited and there is no physical identification of the axis point as there is in a divided detector. General indications are that the best linearity is obtained with single-axis units. For example, data sheets from SiTek Laboratories (Sweden) show typical nonlinearities of ± 0.1 percent for single-axis units and ± 0.5 percent for two-axis units.

The linear CCD has the great advantage that the pixels are fixed in the detector and are in one-to-one correspondence to the angular offset of the returned beam. If pixels are displaced from the ideal location as a result of manufacturing errors, a one-time calibration is sufficient to correct for this. As has already been indicated, it is desirable to spread the image of the source aperture over several pixels in order to obtain a better centroid of the image. A simple approach is a small amount of defocus. If done correctly the centroid can be determined to one-twentieth of a pixel. Since the pixels must be read individually and the resulting signals processed to obtain the centroid, there is an intrinsic delay in the output. This may be significant for applications requiring higher bandwidths and is the chief drawback to the use of CCD detectors in this application.

IV. Evaluation of Commercial Autocollimators

In an effort to find a replacement for the Kollmorgen-built instruments a number of commercially available autocollimators were evaluated. Most were evaluated in the autocollimator test facility at JPL. In two instances the manufacturer's facility was visited, and for all instruments there were detailed technical discussions with the vendor or his representatives. The evaluation criteria were whether the instrument (1) could meet the present specification (Table 1); (2) could meet the optical, electrical, and mechanical interface requirements on the antenna; and (3) would be easy to maintain.

In considering the requirements, only the interface constraints of the intermediate reference structure were taken into account, and the question of operation from the master equatorial was not addressed. This limitation on the scope of the study was adopted because all autocollimators investigated were packaged as an optical head connected to a separate electronics box and were of sizes and shapes which precluded installation on the master equatorial without extensive re-engineering.

In the following discussion autocollimators are grouped by manufacturer. There is no particular significance to the order in which they are discussed.

A. Davidson Optronics

The optical head of the Davidson instrument, model D-696, is physically large. The optical head is approximately $280 \times 230 \times 203$ mm ($11 \times 9 \times 8$ inches) with a lens aperture of 50.8 mm (2 inches) and a lens barrel diameter of 72.7202 mm (2.863 inches). The electronic unit is $610 \times 405 \times 255$ mm ($24 \times 16 \times 10$ inches). It has been in use for many years and is a standard for a number of military applications. The instru-

ment, as was discussed above, uses gas-discharge tubes with a relatively short life. This factor, together with the large physical size and attendant mounting difficulties, led to a decision not to test this instrument.

B. Micro-Radian

Micro-Radian has offered a temperature-compensated instrument based on single-axis optical heads 203 mm (8 inches) long and 50.8 mm (2 inches) in diameter for a number of years. The lens aperture is 25.4 mm (1 inch) and the barrel is 38.1 mm (1.5 inches) in diameter. For two-axis operation, model 145D, two optical heads are mounted side by side and connected to a single electronics unit of approximate dimension 305 × 230 × 155 mm (12 × 9 × 6 inches). Since the dual-axis instrument uses single-axis optical heads, only single-axis instruments were evaluated. Two examples of the single-axis instrument, both model 150, were extensively tested in the JPL facility, although not all tests were performed on both instruments. Transfer curves were measured at various working distances; cross-coupling and NEA were measured; all-attitude sensitivity was checked; and phase shift was measured. Both instruments showed good transfer-curve characteristics at shorter working distances but did not meet requirements at 2.4 m. One instrument showed acceptable linearity at about 1.8 m but did not return enough signal at 2.4 m. The other instrument would operate at 2.4 m but did not meet the linearity specifications in the vicinity of zero.

The phase shift tests were made with a variable-frequency oscillating mirror and were carried out at various settings of the bandwidth control. It was found that at 300 Hz bandwidth the phase shift at 5 Hz was negligible. At 100 Hz bandwidth the phase shift at 5 Hz was approximately 10 degrees.

The double-barreled construction of the optical head used by Micro-Radian prevents it from being a direct substitution for the Kollmorgen 874 currently being used on the antennas, and some modification of the optical assembly on the intermediate reference structure would be required. The simplest modification would be to increase the size of the reflector closest to the autocollimator to accommodate both beams. The most elaborate modification would be to extend the intermediate reference structure, bringing the optical assembly much closer to the master equatorial, and to mount the instrument in looking directly at the master-equatorial mirror, thus minimizing the number of reflections. If such extensive modifications were made, the successful operation of the Micro-Radian instrument could be predicted with a high level of confidence. If, on the other hand, the optical assembly were modified only the minimum amount, the working distance would remain large and the four relay reflections would remain in the optical path. Since the test results indi-

cate marginal performance under these conditions, the Micro-Radian instrument was dropped from further consideration.

C. United Detector Technology

For a number of years UDT has offered a line of position sensors based on quadrant detectors and lateral-effect photodiodes. The line consists of a variety of optical heads for different applications and a number of alternative electronic units for signal processing and data output. The first autocollimator in this line was the model 1000 and was a modular unit with interchangeable lenses and detectors. It was tested for this application with both types of detectors and with both the 200 and 400 mm focal length lenses. The results of the tests were promising but there were shortcomings. Chief among these was the modular construction with joints which would not provide the required long-term mechanical stability. In addition, the best performance was obtained with the 400 mm lens which had an 88.9 mm (3.5 inch) barrel diameter and a 69.85 mm (2.75 inch) aperture. This presented problems with both the optical and mechanical interfaces.

Some time after the completion of the above tests a model 1010 was introduced. The optical head for this was a non-modular, compact unit with a lens barrel of 38.1 mm (1.5 inches) and a 25.4 mm (1 inch) aperture. The small size of this configuration removed the optical and mechanical interface problems, and the one-piece construction removed the potential problems with the modular approach. As soon as an instrument became available it was tested extensively. The results were extremely poor in terms of sensitivity, stability, and NEA. These results were reported to the manufacturer, who undertook an investigation of the manufacturing process. It was finally determined that an incorrectly specified beam splitter had been used in the first batch of instruments.

An instrument with the correct beam splitter was subsequently obtained and tested. At a working distance of 2.1 m (84 inches) one axis could be measured and gave reasonable results. The other axis, however, did not yield useful output. This was traced to poor alignment of the illumination from the LED source with the optical axis of the instrument. A third instrument was obtained and showed the same problem. For this instrument the misalignment was measured to be approximately three-quarters of a degree.

The conclusions were that the model 1010 could only be used if the working distance could be reduced and if the manufacturer were able to improve the alignment of the illumination with the optical axis. The question of thermal sensitivity was not addressed. Even if the other difficulties could be overcome, this would remain a potential problem.

D. Möller-Wedel

The Möller-Wedel autocollimator (J.D. Möller Optische Werk GmbH, West Germany) came on the market late in this study. Its optical head has a 65 mm (2.559 inches) lens barrel and an aperture of 50 mm. The overall dimensions of the head are approximately 380 × 203 × 127 mm. Fairchild 1728 pixel linear CCD array detectors are used for each axis. The centroids are calculated by a combination of analog and digital processing and output digitally on an RS-232 port.

Since the instrument has a digital output it was necessary to design and build an interface unit to meet the electrical interface requirements for analog signals. To allow time for this, the instrument was leased for an extended period rather than borrowed for a single test. The interface unit receives the digital data stream from the autocollimator, separates the data for the two axes, and converts it into two analog voltages of appropriate size. The interface unit is described in more detail in Section V.

First tests indicated that the instrument was extremely precise, had adequate working distance, and generally met all specifications with ample margins, except all-attitude sensitivity and speed of response, which in the original configuration was much too slow. These points were discussed with Wilhelm Duis, the cognizant engineer from Möller-Wedel, during a visit to JPL. He proposed two modifications: (1) a stiffening collar at the junction of the lens barrel and the body of the optical head; and (2) a high-speed data mode in which there was no averaging of successive centroids and no use of the stored calibration. With these modifications the all-attitude specification of three arc seconds peak to peak was met and the data rate raised to 100 readings per second. The removal of the averaging and stored-calibration routines had an insignificant effect on the performance for this application. The linearity remained excellent and the NEA was acceptable.

Since there is a fixed processing and communication delay, the time constant was chosen rather than the phase shift as being a better representation of the instrument characteristics. This measurement was made by introducing a step change in the minor angle and photographing the analog output from the interface unit on an oscilloscope. The time constant was measured to be 0.0134 sec. If a simple RC filter is assumed, this corresponds to a phase shift of 23 degrees at 5 Hz.

Since the decision was made to purchase the instrument and incorporate it into the autocollimator test facility, and since other options presented fewer interface problems, it was decided to defer further work with the Möller-Wedel instrument as a replacement for the Kollmorgen 874. The

work of preparing the required interface electronics was completed and will be described in the next section.

V. Development Work at JPL

In addition to the maintenance of the Kollmorgen instruments and the testing of commercial instruments, three developments have been carried out at JPL. They are (1) the improvement of the Kollmorgen 874 without changing its working principle; (2) the replacement of the scanning system of the Kollmorgen instrument by an array detector without changing the optical or mechanical configuration or the electrical interface; and (3) the construction of an interface for the Möller-Wedel. The motivation for these specific tasks was preparation for the Neptune encounter, and their scheduling and the level of effort were strongly influenced by this fact. Accordingly, development was undertaken in parallel with the plan of dropping the less promising approaches at the appropriate time.

A. Improvements to the Kollmorgen 874

The improvements undertaken in this task leave the basic operating principle of the model 874 unchanged. It remains a mechanically scanned system. The major improvements are in the areas of illumination, mechanical reliability, and circuit stability.

The original design of the model 874 utilized a separately mounted lamphouse connected to the instrument by a bundle of optical fibers. The design depended on the random arrangement of fibers in the bundle to assure uniform illumination over the source aperture. This often resulted in unsatisfactory source uniformity, changes in illumination and the attendant zero shift when the fiber bundle was moved, and problems with damaged fibers. To overcome these problems, a new lamphouse was designed keeping the same transformer and lamp but mounting directly to the instrument housing. The flexible fiber bundle was replaced with a rigid light pipe, and the diffuser directly behind the aperture was improved by shaping the side toward the incident light into a lens. These changes resulted in a significant improvement in light-source performance.

The mechanical scanning in the model 874 is provided by a nutation plate rotating at 9000 rpm. The motor which drives it must operate at approximately 5600 rpm, and motor life has long been a problem. One of the modifications was to install a longer-lived, higher-torque motor. This entailed a new motor mount and an adjustment of the supply voltage.

The approach to the problems of light source variations and the variation of returned flux with distance constitutes a sig-

nificant part of autocollimator design. The approach used in model 874 is that of automatic gain control. The level of the direct current from the detector is used to control the gain of the ac amplification to keep the sensitivity to angular variations (slope of the transfer curve) constant. The experience with the AGC of the 874 has been quite varied. For some units the AGC works well and a constant output can be maintained for wide variations in the flux on the detector. For other units there have been problems with excessive phase shift and instabilities if the circuit is adjusted for good AGC operation.

Since the Mark IVA firmware can accommodate changes in the slope of the transfer curve and since the transfer curve slope can be measured in the field by fixing either the antenna or the master equatorial and scanning the other, it was decided to replace the AGC circuit with a simple gain control. Under this approach the instruments are shipped out with a gain setting which gives the specified slope to the transfer curve at 2.44 m working distance in the laboratory environment. When the instrument is installed for the antenna, the gain is measured and entered into the control computer, and this gain is checked from time to time as required.

The changes outlined here as well as some other minor circuit improvements are being implemented in instruments as they come in for routine repair and recalibration.

B. Replacement of Scanning System by a Circular Detector Array

To this point the only options that have been discussed have been upgrading the model 874 without changing the operating principle or replacing it completely with a commercial instrument. A third option is also being pursued: that of keeping the optical and mechanical configuration of the 874 intact but replacing the scanning system with an array detector. This has the considerable advantage of leaving the optical and mechanical interfaces unchanged. The approach had been under consideration for several years, but lack of a suitable detector had blocked progress. When the Parkes Radio Astronomy Observatory in Australia reported good results with a Reticon RO-64 circular array detector, it was decided to investigate its use in the model 874. The chief issues were those of sufficient photometric response and sensitivity to variations in illumination across the source aperture. Promising results from simulation studies and laboratory test led to a prototype test on DSS-14 in September 1987. Except for a minor problem with the acquisition indicator, the performance was satisfactory during a month of testing. The decision was made to proceed with an engineering model, which is now under construction.

The operating principle of the modified instrument (designated 874-64) is very similar to that of the original model 874

except that the scanning of the array has replaced the mechanical scanning. The RO-64 detector consists of 64 silicon photodiodes in a circular array 2 mm in diameter. In response to a clock signal the elements are read out continuously in sequence. This has an advantage, in addition to the quite obvious elimination of mechanical scanning, in that each of the 64 elements integrates continuously over the time interval between readouts. As a consequence, signal-to-noise ratios are greatly improved over those of a scanning system of comparable resolution. Furthermore, scanning rates may be much higher than for mechanical systems, resulting in higher bandwidth systems. A circular source aperture is used, and the image on the detector is slightly defocused to produce tapering illumination at the edge of the image. When the image is centered on the array, all pixels in the array give the same signal. As the center of the image moves in a particular direction, the signals from the pixels toward which the image is moving increase and those diametrically opposite decrease.

The block diagram for the electronics is shown in Fig. 7. The clock, counters, and dividers generate the clocking signals for the circular array and the reference signals for the demodulator. A synchronizing signal generated by the array on the completion of each scan synchronizes the demodulation signals with position on the array and enables the demodulators to output signals corresponding to designated X and Y directions. A by-product of this is that it is possible to electronically "rotate" the array to place the axes in the desired directions. The signal from the detector consists of a train of pulses at a constant repetition rate amplitude modulated by the variation of the light intensity around the array. Although the functions are not in fact sequential, the demodulation and filter circuits may be thought of as performing three functions: (1) detecting the low frequency modulation on the pulse train; (2) synchronously rectifying this modulation in each of two channels; and (3) filtering the rectified signal to produce dc signals proportional to the offset.

The model 874-64 is able to meet the specifications in all areas except possibly acquisition range. However, such a large acquisition range is not required operationally. The phase shift at 5 Hz is well below the specified values. It was estimated to be 0.2 degree, which is at the limit of what can be measured.

C. Möller-Wedel Interface

It has already been mentioned that it was necessary to construct an interface unit to permit the Möller-Wedel autocollimator to be used as a replacement for the Kollmorgen 874. The instrument as modified for this particular application had two data modes at the RS-232 port: a normal mode using 2400 baud ASCII characters and a high-speed rate, 9600 baud

absolute binary. It is the high-speed rate that must be processed for control applications. The basic operation in the interface unit consists of reading the digital output of the Möller-Wedel, applying zero-point and scale-factor corrections, converting to analog voltages, and outputting the properly scaled voltage for each axis. In addition the unit must test for good data and close an acquisition relay when valid data is being processed.

The interface unit is built around a 990/101M computer based on Texas Instruments' TMS9900 16 bit microprocessor. It has four kilobytes of RAM, four kilobytes of PROM, two kilobytes of non-volatile CMOS memory (battery backup), and two RS-232 ports. The D/A converters are memory mapped I/O ports and are followed by gain stages to achieve the required 64 V/degree sensitivity. Both the Möller-Wedel and the interface unit operate from the 60 Hz line, but power is switched by the 400 Hz line that operates the 874 autocollimators. It was necessary to make a small modification to the Möller-Wedel to make it power up in the high-speed data mode. With these modifications, the Möller-Wedel will resemble the 874 in all operational aspects.

There are three programmed modes of operation for the interface unit. They differ in the values taken for the scale zero. The first program uses the factory calibrated zero. The second takes the first value read and calls that "zero." The third program uses the constants stored by the second program for all subsequent measurements, and is the one used for the normal control function.

In mounting the Möller-Wedel autocollimator on the antenna the optical head would mount in place of the 874. The electronics box would be mounted close to it on the IRS. The interface unit would be on the floor of the master-equatorial

room and would connect to the junction box on the wall of the room rather than to the one on the IRS.

VI. Summary and Recommendations

A number of commercial autocollimators have been evaluated. Based on the assumption that modifications to the intermediate reference structure would be minimal, no instruments with analog output were judged satisfactory as a replacement for the Kollmorgen 874. A digital instrument made by Möller-Wedel, after factory modification, showed sufficient promise that an interface unit was designed and built. It has not been tested on the antenna because other solutions developed in parallel offered promise of simpler solutions for the Voyager Neptune encounter period.

For the Neptune encounter a two-step approach has been taken to the upgrading of the Kollmorgen 874. The first step seeks to improve the reliability without changing the basic technology of the instrument. The changes include a new illumination system, a new motor, and the replacement of the AGC circuit with a manual gain control. These changes are being implemented as instruments come in for routine maintenance and recalibration. The second step involves the more significant change of replacing the mechanical scanning with a circular array detector. A prototype has been tested at DSS-14, and an engineering model is now under construction. Following its testing the decision will be made concerning the conversion of other units.

For the period after the Neptune encounter, careful consideration should be given to digital-output instruments. Möller-Wedel has under development a high-speed, single-axis autocollimator with parallel digital output. It uses a high-intensity LED and operates at a 400 Hz update rate. The evaluation of this instrument is strongly recommended.

References

- [1] T. H. Thurston, "Electronic Autocollimators," in *Proc. SPIE*, vol. 251, pp. 46-50, 1980.
- [2] L. D. Hutchenson, "Practical Electro-Optic Deflection Measurement System," *Optical Engineering*, vol. 15, pp. 61-63, 1976.

Table 1. Principal specifications for the autocollimator for the 70-meter antennas

Slope of the transfer curve (V/degree)	64
Linear range (arc seconds)	±300
Linearity (percent)	±10
Accuracy at null (arc seconds)	±1
Acquisition range (arc minutes)	±23
Maximum working distance (meters)	2.44
RMS noise-equivalent angle (arc seconds)	0.5
Maximum phase shift 0–5 Hz (degrees)	5
Cross-coupling (percent)	±2
All-attitude sensitivity (peak to peak, arc seconds)	3

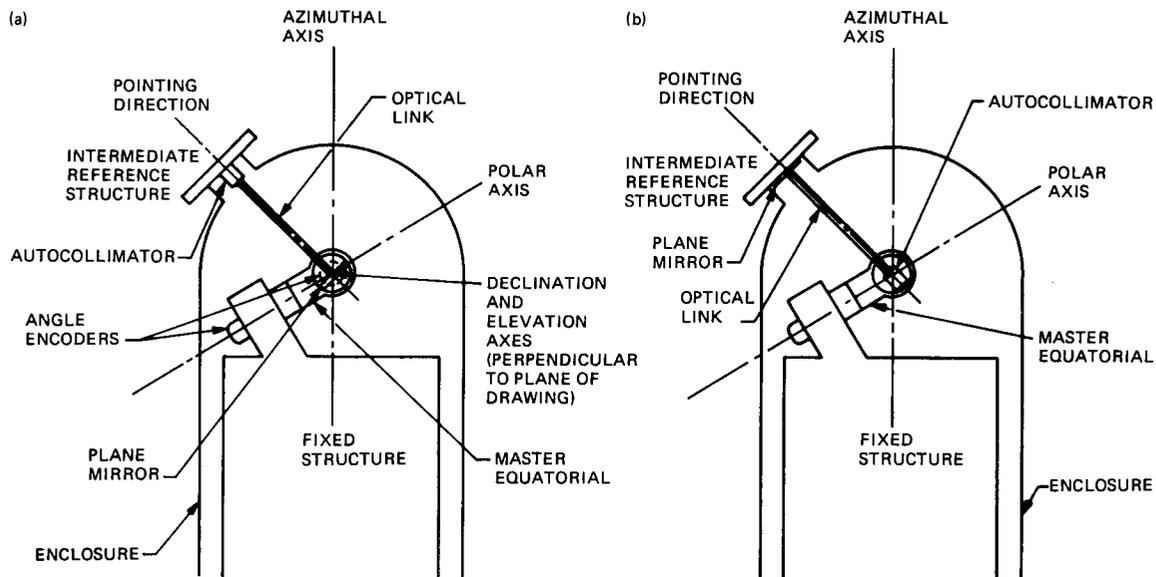


Fig. 1. Pointing-metrology system for the 70-meter antennas. In (a) the antenna follows the master equatorial; in (b) the master equatorial follows the antenna.

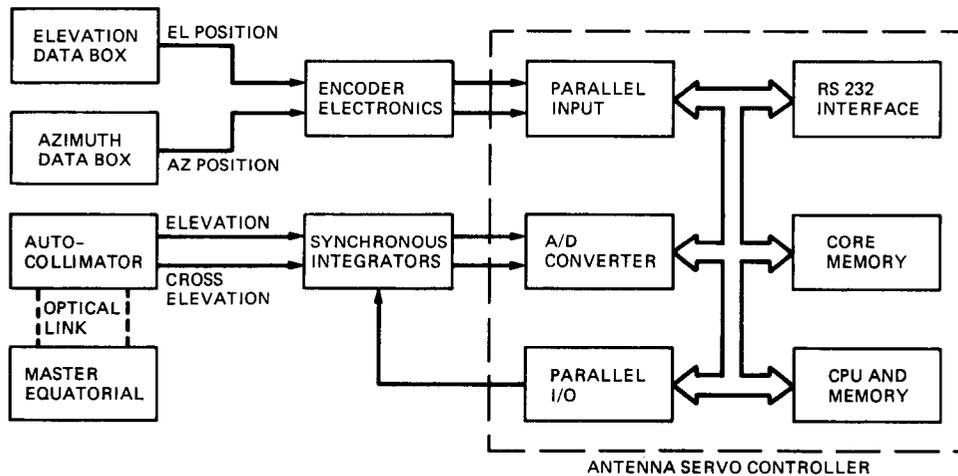


Fig. 2. Block diagram showing signal flow from the autocollimator to the antenna control system

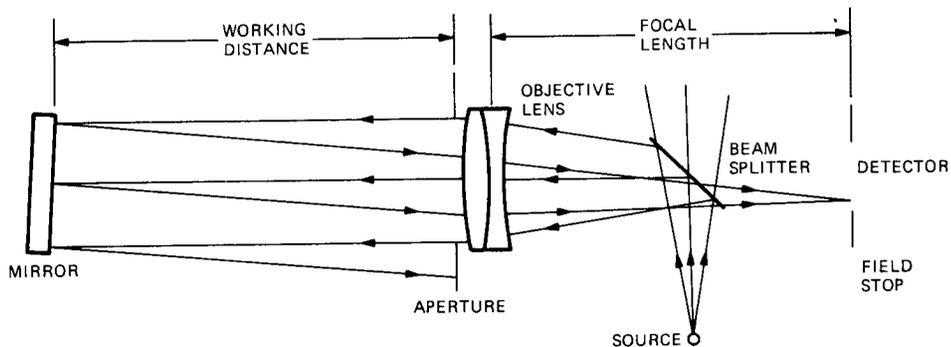


Fig. 3. General autocollimator configuration. The source and detector may be interchanged.

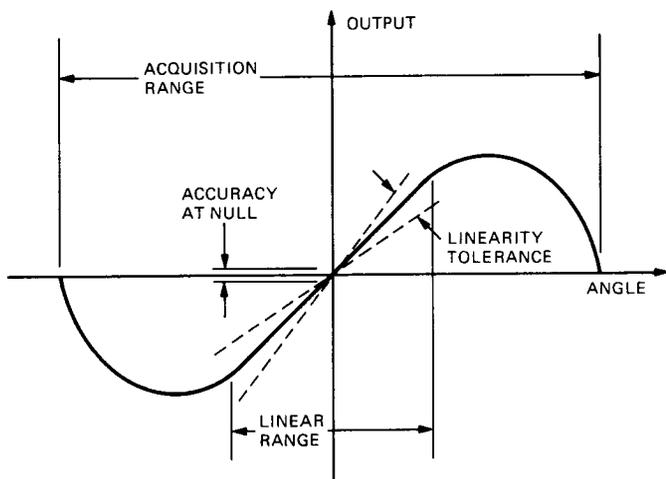


Fig. 4. Typical autocollimator transfer curve showing the relationships among several parameters

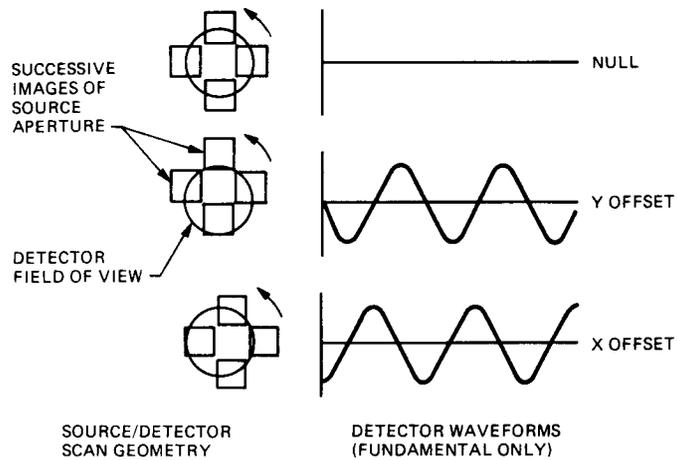


Fig. 5. Mechanical scanning used in the Kollmorgen model 874 autocollimator. The waveforms for representative offsets are shown.

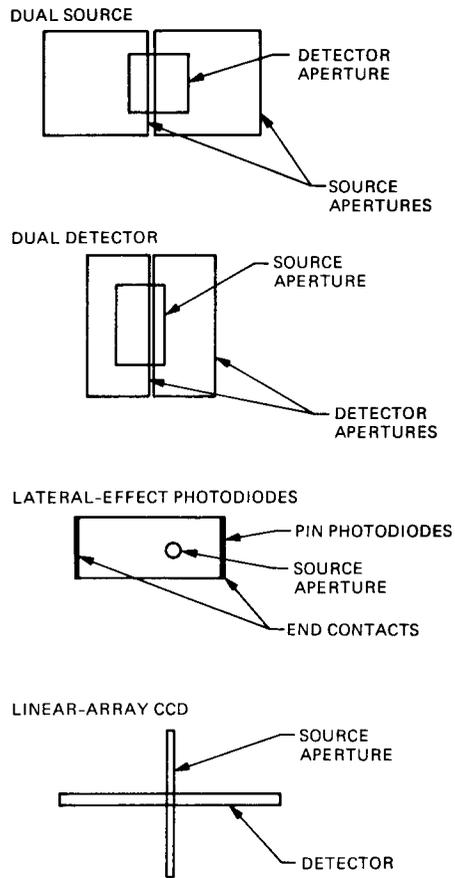


Fig. 6. Non-scanning detection schemes used in commercial autocollimator

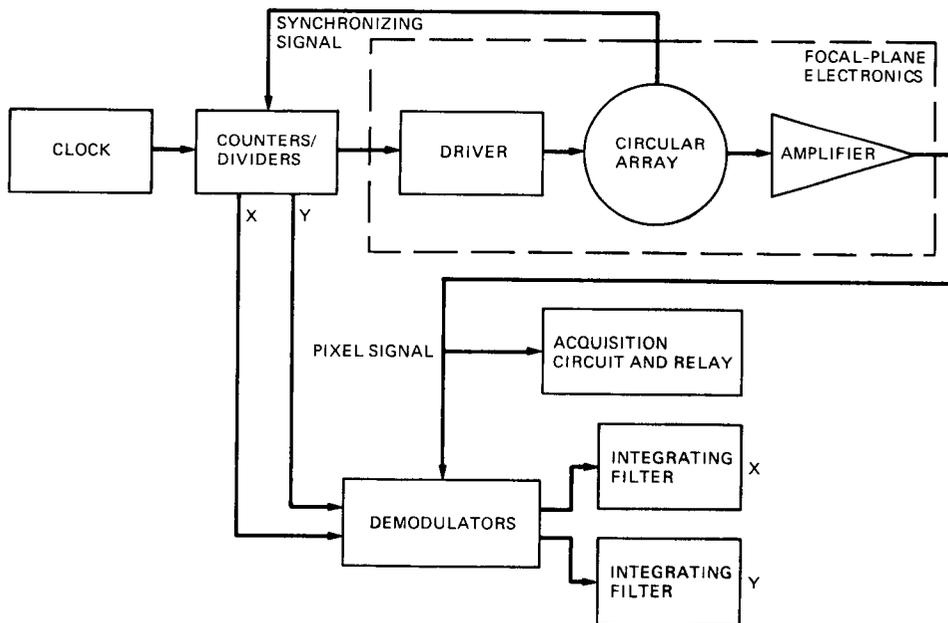


Fig. 7. Block diagram of the circular-array detector implementation in model 874-64

519-32

123-35

TDA Progress Report 42-92

AS

N88-18793

October-December 1987

Seventy-Meter Antenna Performance Predictions: GTD Analysis Compared With Traditional Ray-Tracing Methods

J. M. Schredder*

Ground Antenna and Facilities Engineering Section

A comparative analysis has been performed, using both the Geometrical Theory of Diffraction (GTD) and traditional pathlength error analysis techniques, for predicting RF antenna gain performance and pointing corrections. The NASA/JPL 70-meter antenna with its shaped surface was analyzed for gravity loading over the range of elevation angles. Also analyzed were the effects of lateral and axial displacements of the subreflector. Significant differences were noted between the predictions of the two methods, in the effect of subreflector displacements, and in the optimal subreflector positions to focus a gravity-deformed main reflector. The results are of relevance to future design procedure.

I. Introduction

Among a number of current trends in high performance antenna design is the replacement of paraboloid main reflectors and hyperboloid subreflectors with optimally shaped surfaces which provide uniform aperture illumination. Quantification of various RF gain loss mechanisms, especially those due to surface imperfections, is essential to understanding where cost-effective improvements might be realized. The traditional methods used for antenna gain and pointing analysis have been based on the assumption of paraboloid main reflectors and hyperboloid subreflectors. Newer methods

exist which deal with more general reflector shapes. One such method is implemented in the JPL Geometric Theory of Diffraction (GTD) program. This article compares gain and pointing predictions derived from traditional methods and GTD analysis under various conditions. The NASA/JPL 70-meter antenna, which has a shaped main reflector and a shaped subreflector, was used as a test sample. The effect of lateral and axial offsets of the subreflector was investigated along with the effect of gravity deformations of the main reflector with focusing of the subreflector.

II. Traditional Ray-Tracing Methods

For the analysis of gravity deformations of the main reflector, the traditional and GTD methods use the same set of nodal displacements, derived from a finite-element struc-

*Mr. Schredder, who is assigned on contract to the Ground Antenna and Facilities Engineering Section, is an employee of Planning Research Corporation, McLean, Virginia.

tural model. These are the X , Y , and Z (axial) components of symmetric unit-gravity-load deflections ($\Delta S_{\text{symmetric}}$) and antisymmetric unit-gravity-load deflections ($\Delta S_{\text{antisymmetric}}$). The two sets of deflections are combined according to the following equation:

$$\begin{aligned} \Delta S = & \Delta S_{\text{symmetric}} [\sin(EL_{\text{rig}}) - \sin(EL)] \\ & + \Delta S_{\text{antisymmetric}} [\cos(EL_{\text{rig}}) - \cos(EL)] \end{aligned} \quad (1)$$

where:

$$S = X, Y, \text{ or } Z$$

to yield the deflections at any desired elevation angle (EL). The rigging elevation angle (EL_{rig}) was arbitrarily set at 45 degrees.

In ray-tracing analysis, the half-pathlength errors are then fitted to a paraboloid by least squares, and the RMS of residuals of the fit is computed [1]. The deformed antenna is assumed to be perfectly focused. The RMS value is used in the Ruze equation [2] to compute the gain loss as follows:

$$\zeta = e \exp \left[- \left(\frac{4\pi(\text{RMS})}{\lambda} \right)^2 \right] \quad (2)$$

where:

$$\zeta = \text{efficiency}$$

$$\lambda = \text{wavelength}$$

The elevation pointing shift due to gravity-loading main reflector and quadripod deflections is given by:

$$\theta = \theta_s [\sin(EL_{\text{rig}}) - \sin(EL)] + \theta_{as} [\cos(EL_{\text{rig}}) - \cos(EL)] \quad (3)$$

where θ_s is the pointing shift due to symmetrical gravity loading and θ_{as} is the pointing shift due to antisymmetric loading. For the 70-m antenna, $\theta_s = 0.0015$ degree and $\theta_{as} = 0.047$ degree.

This equation is derived from the finite element model and the best-fit paraboloid axis tilt. The fitting process also gives the parameters of the best-fit paraboloid, including its focus. The position of the subreflector vertex, the subreflector axis tilt, and the feedhorn phase center shift, as obtained from the whole antenna structure model, are combined with the main reflector best-fit geometry. A subreflector focusing offset

table is generated which brings the virtual image of the feed phase center into coincidence with the best fit main reflector focus [3]. This is shown in Fig. 1. The pointing shift predicted by Eq. (3) is added to the shift computed by Eq. (4) to derive a predicted net shift for the focused antenna.

Gain loss resulting from subreflector offsets is computed traditionally as follows. Data obtained from running the JPL Radiation Program [4] allowed the equivalent RMS pathlength error per unit subreflector displacement in lateral and axial directions to be expressed as functions of focal length to diameter ratio. For the shaped 70-meter antenna, the approximating f/D ratio is taken to be 0.389, which gives an RMS pathlength error of 0.0773 cm per centimeter of axial displacement and 0.0185 cm per centimeter of lateral displacement. These values are then used in the Ruze equation (Eq. [2]).

Pointing shift as a function of subreflector lateral displacement for a Cassegrain antenna is predicted by a simple geometric argument, given in [5]:

$$\theta = -\Delta Y \left(1 - \frac{1}{M} \right) \frac{K}{f} \quad (4)$$

where:

$$\Delta Y = \text{lateral subreflector displacement}$$

$$\theta = \text{pointing shift in radians}$$

$$f = \text{focal length}$$

$$K = \text{beam deviation factor}$$

$$M = \text{magnification factor}$$

For the 70-m antenna, $f = 2722.9$ cm, M is estimated as 6.84, and K is estimated as 0.82.

III. The GTD Method

The GTD program evolved from an electric field integration program developed in 1978 and modified in 1983.¹ The program uses modified Jacobi polynomials to describe the radial dependence of the surface currents induced by the fields incident on the main reflector. The modified Jacobi polynomials are an orthogonal set with desirable convergence properties. The program uses two-dimensional Gauss integra-

¹Y. Rahmat-Samii, "Offset Parabolic Reflector Computer Program for Analysis of Satellite Communications Antennas," JPL Publication D-1203 (internal document), Jet Propulsion Laboratory, Pasadena, California, December 1983.

tion to determine the coefficients for the eigenfunctions composed of the products of modified Jacobi polynomials and sine and cosine functions of the azimuth angle. These coefficients are then used in another series to compute the far-field electric field pattern.

The program described in JPL Publication D-1203¹ requires that the user supply subroutines to compute the following: a description of the main reflector surface that includes axial distance (Z) as a function of radius (r) and azimuth angle (ϕ), the first derivatives $\partial z/\partial r$, $\partial z/\partial \phi$, and the fields incident on the main reflector.

The GTD program represents the deformed or undeformed main reflector as a series of modified Jacobi polynomials added to a base paraboloid:

$$Z = \sum_n \sum_m (C_{nm} \cos n\phi + D_{nm} \sin n\phi) F_m^n \left(\frac{r}{R_{\max}} \right) + z_0 + \frac{r^2}{4f} \quad (5)$$

where:

ϕ = azimuth angle

r = radius

R_{\max} = the radius of the main reflector

f = focal length of base paraboloid

z_0 = arbitrary datum

The program computes two sets of electric and magnetic fields incident on the main reflector: a pair of fields reflected from the subreflector and a pair of fields diffracted from the subreflector. The reflected fields are computed using geometrical optics (GO), while the diffraction field is computed using the Geometrical Theory of Diffraction (GTD). The program in this form is documented in JPL Publication D-2583.²

IV. GTD Analysis of the 70-Meter Antenna

In this investigation, a number of modifications were made to the GTD program to accommodate the 70-meter antenna analysis. An additional data block was generated to provide storage for Jacobi polynomial coefficients for both the shaped subreflector and the deformed main reflector (previously, the

program could analyze one or the other, but not both). Also, a change was made in the sequence in which Jacobi polynomial values were computed and stored and provided a major increase in execution speed.

Analysis of a gravity-deformed reflector involves the following five steps:

- (1) The computation of node deflections at a given elevation angle is the same as in the conventional analysis, and the deflections (u , v , w) are added to the node coordinates (x , y , z) for an undeflected reflector to arrive at a set of deflected nodes. The 70-meter model that was used has 764 nodes for a half-model.
- (2) A grid of axial positions $z(r, \phi)$ at evenly spaced values of radius (r) and azimuth angle (ϕ) is generated. The radius varies from zero at the center to R_{\max} , the radius of the main reflector, and the azimuth angle varies from 0 to 360 degrees. To generate the grid, a set of nine neighboring nodes around each grid point is used, as shown in Fig. 2. First, three interpolating parabolas along points 1-2-3, 4-5-6, and 7-8-9 are computed. These parabolas are evaluated at the grid point radius to give three values of azimuth angle and Z at this radius. Then a fourth interpolating parabola is computed giving Z as a function of angle. This parabola is evaluated at the grid point angle to yield the interpolated Z value. Typical grid spacing was 416 intervals in radius and 256 intervals in angle.
- (3) The following equation is integrated by standard numerical methods to generate the Jacobi polynomial coefficients (C_{nm} , D_{nm}) which describe the reflector surface:

$$C_{nm} = \frac{\epsilon_n}{2\pi} \int_0^{2\pi} \int_0^1 \cos n\phi F_m^n(s) s \, d\phi \, ds \quad (6)$$

$$D_{nm} = \frac{\epsilon_n}{2\pi} \int_0^{2\pi} \int_0^1 \sin n\phi F_m^n(s) s \, d\phi \, ds$$

where:

$$s = r/R_{\max}$$

$$\epsilon_n = 1 \text{ if } n = 0$$

$$\epsilon_n = 2 \text{ if } n \neq 0$$

To represent the gravity-deformed main reflector shapes, a 7×25 set of coefficients was used ($n = 0$,

²T. Veruttipong, *et al.*, "Dual Shaped and Conic GTD/Jacobi-Bessel Analysis Programs," JPL Publication D-2583 (internal document), Jet Propulsion Laboratory, Pasadena, California, July 30, 1985.

1, . . . , 6; $m = 0, 1, . . . , 24$), which can model 24 ripples in radius and 6 cycles in azimuth angle.

- (4) To check the accuracy of the Jacobi polynomial fit, the reflector surface is reconstructed from the Jacobi polynomial representation (at the node radii and angles). To do this, the coefficients determined in step 3 are entered into Eq. (5). The computed Z values are compared with the Z values derived from the finite element model. For gravity deformations, RMS differences of 0.075 mm (0.003 inch) to 0.175 mm (0.005 inch) were found. These differences were much smaller than the gravity deformations.
- (5) The coefficients determined in step 3 are also used to represent the deformed main reflector in the GTD program.

Figure 3 shows the sequence of computations of both the traditional and GTD methods.

The Jacobi polynomial representation of the undeflected main reflector consists of a set of 15 coefficients ($n = 0; m = 0, 1, . . . , 14$) of polynomials in radius, as there is no angular dependence of the ideal reflector surface. The first 15 Jacobi polynomials were sufficient to describe the deviation of the radial profile from a parabola. This representation of the "perfect" reflector was used in two ways: (1) to generate the undeflected node positions for the gravity deformation analysis; and (2) to study the effects of subreflector displacements.

Among the input variables to the GTD program are the subreflector and feed positions; the orientations of the main reflector, subreflector, and feed coordinate systems; and the RF wavelength. The frequency used in this study was 8.45 GHz, which has a wavelength of 35.48 mm (1.397 in.). The study was performed with the subreflector pointed at the X-band horn, as shown in Figs. 4 and 5. The geometry is shown in more detail in JPL Publication D-1843.³

V. Results

The following results are compared between GTD and traditional analysis:

- (1) Gain loss resulting from subreflector lateral and axial offsets, and pointing shift due to lateral offsets.
- (2) Gain loss and pointing shift as functions of elevation angle with the subreflector focused.
- (3) Prediction of best subreflector offsets to focus a gravity-deformed antenna as a function of elevation angle.

It should be noted that the subreflector offsets are given in units of the wavelength at 8.45 GHz.

Figure 6 shows gain loss for a perfect main reflector as a function of axial subreflector displacement predicted by the two methods. For positive axial displacements (away from the main reflector) the agreement is very close, but the difference between the two curves is sizable for negative displacements.

Table 1 shows the gain loss predicted by the two methods for lateral subreflector displacements, while Table 2 shows the pointing shift predicted. There is a large difference in the predicted gain loss; however, both methods predict a square-law dependence of gain loss on lateral subreflector displacement. Also, the pointing shift predicted is somewhat different. Ray-tracing methods predict a shift of 0.01472 degree per centimeter of lateral displacement, while GTD predicts 0.01346 degree per centimeter of lateral displacement.

Figures 7 and 8 show the predicted gain loss and pointing shift, respectively, of the 70-meter antenna with focused subreflector as functions of elevation angle. For each figure, three sets of curves were generated: (1) predictions from traditional methods; (2) predictions from GTD analysis using the subreflector focusing tables furnished by traditional methods; and (3) predictions from GTD analysis with subreflector position varied to maximize the predicted gain. If the gain is maximized by varying the subreflector position, the gain loss predictions of GTD analysis agree with those of ray-tracing analysis to within a few hundredths of a decibel. This is considered good agreement.

Figure 9 contains four curves. The two broken lines represent the subreflector offsets required to bring the virtual image of the feed phase center into coincidence with the focus of the best-fit paraboloid, while the two solid lines represent the results of searching for the subreflector positions which maximize the gain predicted by GTD analysis. Note that in both cases, the offsets are measured from the original position of the subreflector, in the main reflector coordinate system. The agreement appears to be good in Z and poor in Y .

³A. G. Cha and W. A. Imbriale, "Computer Programs for the Synthesis and Interpolation of 70-m Antenna Reflector Surfaces," JPL Publication D-1843 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 1984.

VI. Conclusion

The predictions of traditional ray tracing and GTD analysis have been compared in this article for subreflector displacements and for the focused, gravity-deformed 70-meter antenna. There is a significant difference in the gain loss predicted by the two methods for axial subreflector displacement, and a large difference in the gain loss predicted for lateral displacement. The pointing shift predicted for lateral displacement is also somewhat different.

For the focused gravity-deformed antenna, the gain loss predictions of the two methods show good agreement if the subreflector position is varied to maximize the gain. It is

noteworthy that the pointing shift predictions show close agreement between ray tracing and GTD analysis if the traditional subreflector focusing method is used. However, if the traditional method of determining subreflector focusing offsets is used, the agreement for gain loss predictions is poor. This study shows significant differences between the subreflector offsets that align the virtual image of the feed phase center with the best-fit paraboloid focus and the offsets which maximize the gain. Past studies have indicated that aligning the main focus and feed phase center will yield good results in maximizing the gain of paraboloid-hyperboloid systems. The present results indicate that methods which compute the electric and magnetic fields are required to give good results when dealing with shaped surfaces.

References

- [1] M. S. Katow and L. W. Schmele, *Antenna Structures: Evaluation Techniques of Reflector Distortions*, JPL Space Programs Summary No. 37-40, vol. IV, Jet Propulsion Laboratory, Pasadena, California, August 31, 1966.
- [2] J. Ruze, "Physical Limitations on Antenna," Technical Report 248, Cambridge, Massachusetts: Research Laboratory of Electronics, Massachusetts Institute of Technology, October 1952.
- [3] R. D. Hughes and M. S. Katow, "Subreflector Focusing Techniques Applied to New DSS-15 and DSS-45 34-Meter Antennas," *TDA Progress Report 42-80*, vol. October-December 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 832-890, February 15, 1984.
- [4] M. S. Katow, "Thirty-four-Meter Antenna-Subreflector Translations to Maximize RF Gain," *DSN Progress Report 42-62*, vol. January-February 1981, Jet Propulsion Laboratory, Pasadena, California, pp. 112-120, April 15, 1981.
- [5] R. Levy, "Optimization of Antenna Structure Design," presented at the Eighth ASCE Conference on Electronic Computation, Houston, Texas, February 1983.

Table 1. Comparison of predictions of gain loss caused by subreflector lateral displacements

Lateral displacement	Gain loss, dB (traditional)	Gain loss, dB (GTD)
λ^* in +Y direction	0.23	0.55
$\frac{\lambda}{2}$ in +Y direction	0.06	0.14
λ in -Y direction	-	0.56
λ in +X direction	-	0.56

* $\lambda = 3.548$ cm (1.397 in.) at 8.45 GHz.

Table 2. Comparison of predictions of pointing shift caused by subreflector lateral displacements

Lateral displacement	Pointing shift (traditional), deg	Pointing shift (GTD), deg
λ^* in +Y direction	0.05228	0.04776
$\frac{\lambda}{2}$ in +Y direction	-	0.02388
λ in -Y direction	-	0.04775
λ in +X direction	-	0.04775

* $\lambda = 3.548$ cm (1.397 in.) at 8.45 GHz.

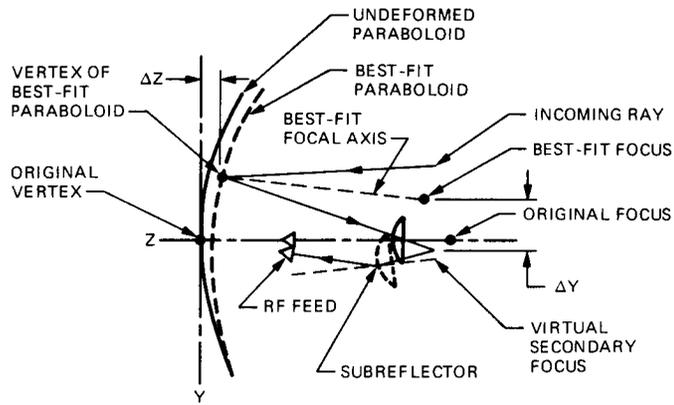


Fig. 1. RF center ray tracing and hyperboloid offsets with gravity distortions

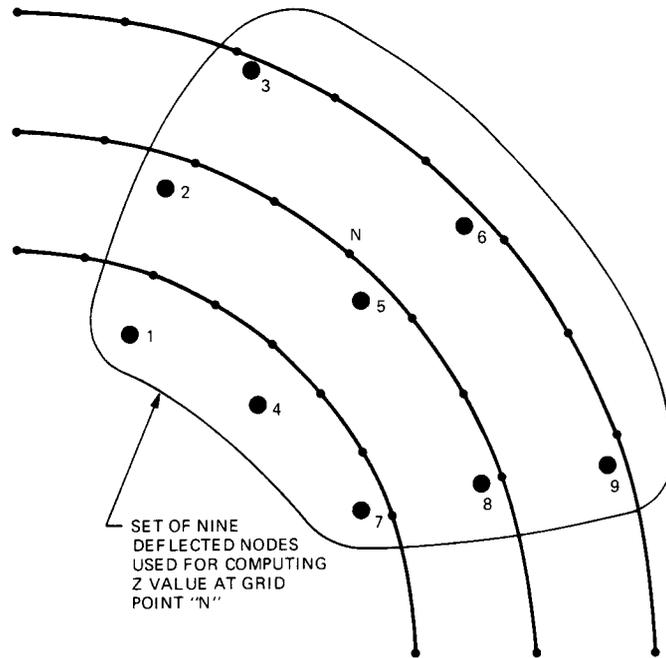


Fig. 2. Use of deflected node coordinates to compute axial (Z) coordinate of a grid point

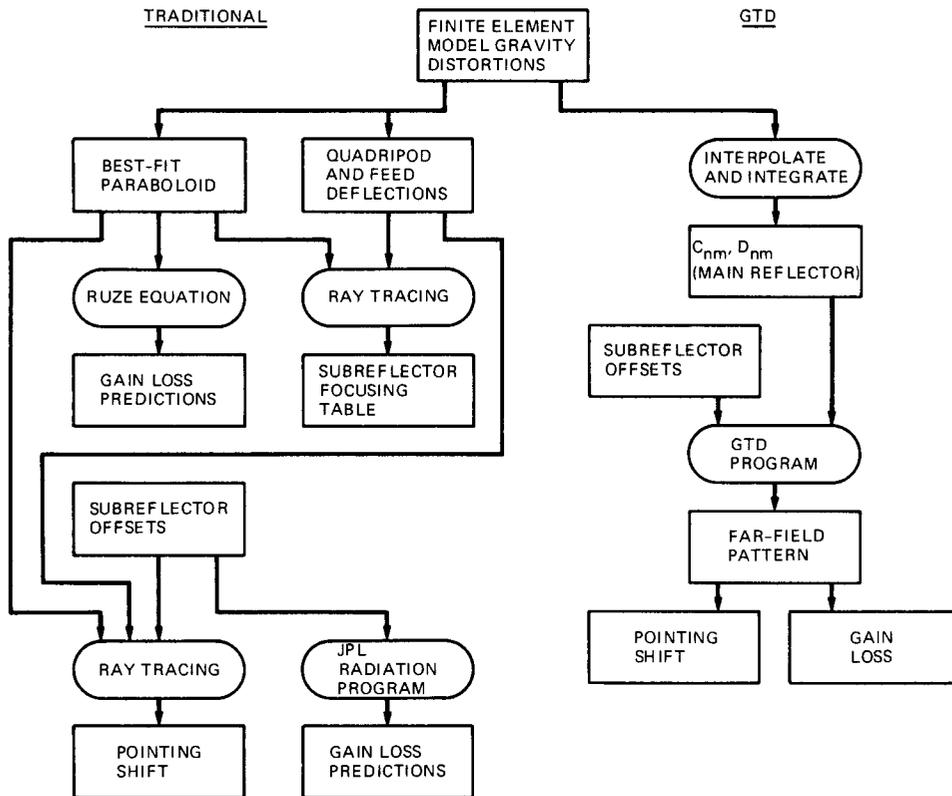


Fig. 3. Sequence of computations for calculating effects of gravity deformations and subreflector offsets by traditional and GTD methods

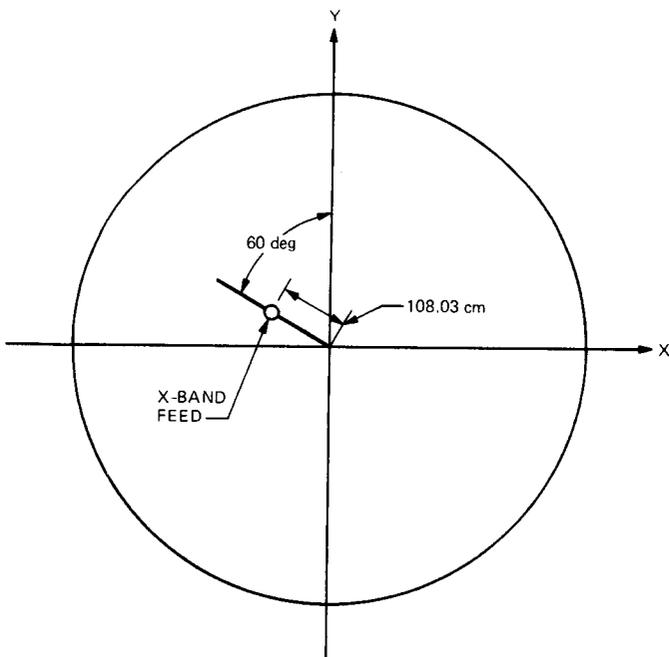


Fig. 4. Antenna geometry used in this study, looking into main reflector from space

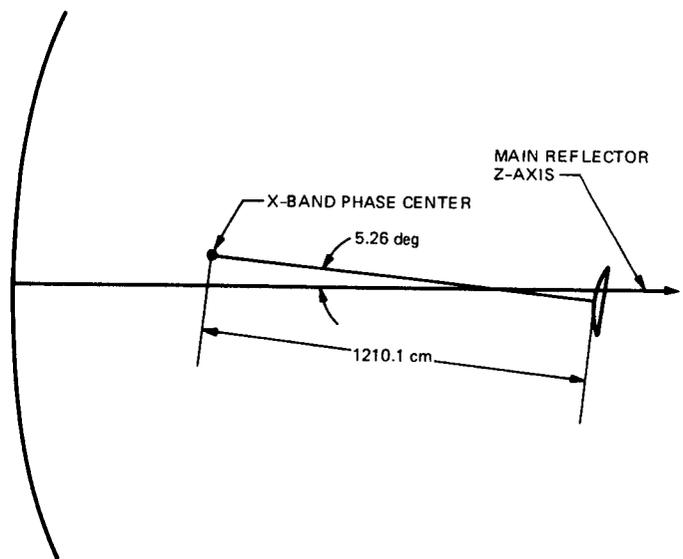


Fig. 5. Antenna geometry, plane containing the feed phase center and the main reflector Z axis

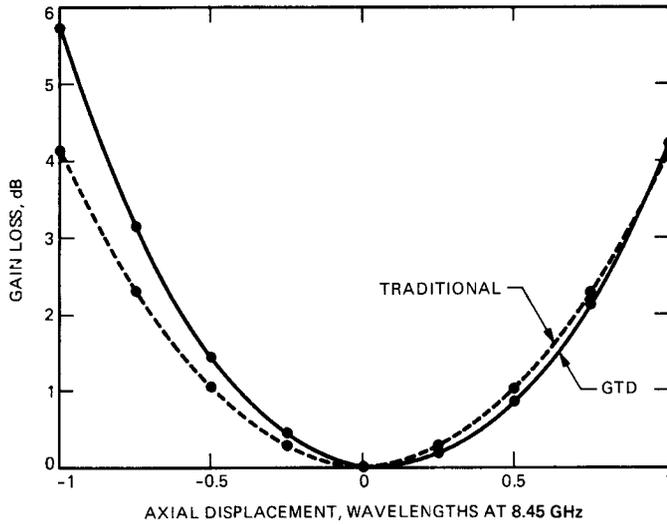


Fig. 6. Gain loss versus axial subreflector displacement

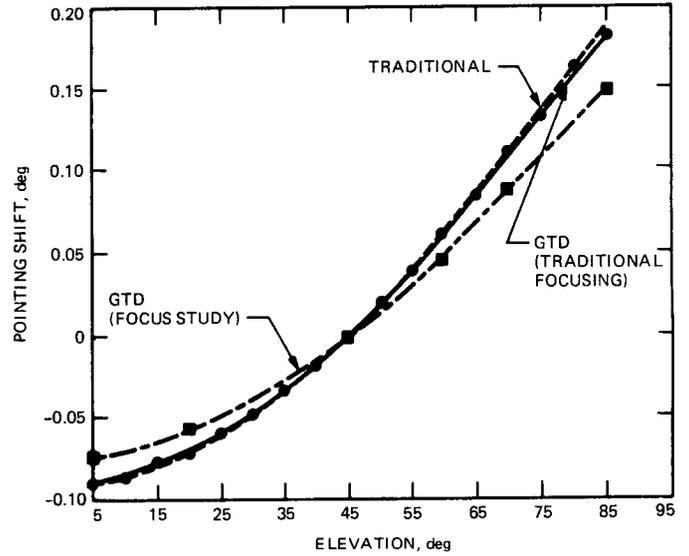


Fig. 8. Elevation pointing shift versus elevation for gravity-deformed, focused antenna

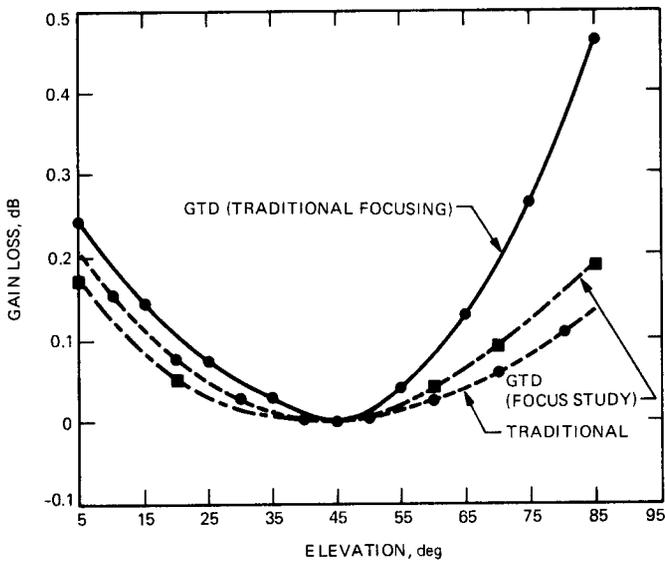


Fig. 7. Gain loss versus elevation for gravity-deformed, focused antenna

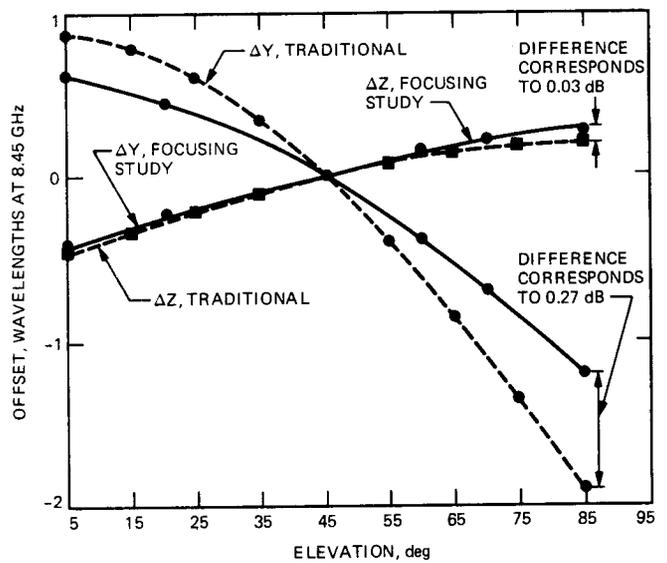


Fig. 9. Subreflector lateral and axial offsets to focus a gravity-deformed 70-meter antenna

The Design and Implementation of the Technical Facilities Controller (TFC) for the Goldstone Deep Space Communications Complex

D. A. Killian and F. J. Menninger
Ground Antenna and Facilities Engineering Section

T. Gorman and P. Glenn
TDA Mission Support and DSN Operations

The Technical Facilities Controller is a microprocessor-based energy management system that is to be implemented in the Deep Space Network facilities. This system is used in conjunction with facilities equipment at each of the complexes in the operation and maintenance of air-conditioning equipment, power generation equipment, power distribution equipment, and other primary facilities equipment. The implementation of the Technical Facilities Controller has been completed at the Goldstone Deep Space Communications Complex and is now operational. This article describes the installation completed at the Goldstone Complex and evaluates the utilization of the Technical Facilities Controller. The findings will be used in the decision to implement a similar system at the overseas complexes at Canberra, Australia, and Madrid, Spain.

I. Background

The Deep Space Network (DSN) is operated and managed for NASA by JPL and is composed of three Deep Space Communications Complexes (DSCCs) located at Goldstone, California; Madrid, Spain; and Canberra, Australia. The DSN serves as the primary facility for communication with deep space missions such as Voyager 1 and 2, the Pioneer series, and the soon to be launched Galileo spacecraft.

The Goldstone Complex consists of four Deep Space Stations (DSSs) extending over a sixteen mile stretch of road in the Mojave Desert. It includes 34-m and 70-m class antennas and about one hundred antenna support buildings ranging in function from control room buildings and generator power

plant buildings to administrative buildings that support approximately 200 employees.

The Technical Facilities Controller (TFC) was initially conceived circa 1975 as a Utility Control System (UCS) when a distributed process controller was developed to monitor and control facilities equipment for energy management. The facilities equipment includes Heating, Ventilating, and Air Conditioning (HVAC), lighting, power generation, power distribution, and site protection equipment.

The initial prototype system was demonstrated at the Venus station (DSS-13) at Goldstone using a scaled-down version of the TFC called "Pathfinder." Pathfinder was envisioned to

monitor and control room temperatures and power. Therefore, it would control energy loads such as HVAC equipment and lighting. The results of this prototype showed that a distributed process controller should be implemented as a useful facilities energy management tool. The energy management features were augmented by many additional advantages such as improving facility performance, reporting capabilities, and instrumentation monitoring.

II. Functional Requirements

The functional requirements for the TFC centered on many operational needs for upgrading facility performance in unattended operation, safety, maintainability, and availability. The following is a fundamental list of functional requirements for the TFC:

- (1) The TFC shall be capable of receiving input data from different sources such as digital sensors (switch closures), analog sensors (4–20 mA signal), real time clock information, and electrical power meters (with pulsers).
- (2) The TFC shall be capable of sending digital on/off load control commands to any TFC connected equipment.
- (3) The TFC shall be capable of accepting input sensor data and sending output load commands at multiple locations distributed throughout the Complex (i.e., serve as a distributed process control and monitor).
- (4) The TFC shall be capable of handling single "event control." An event control is defined as a specific set of input sensor data used to trigger a command to set a specific output load configuration.
- (5) The TFC shall be capable of executing energy management control of specific loads. This control consists of peak demand load shedding during periods of high and costly energy usage, and exercise time-of-day control.
- (6) The TFC shall be capable of archiving sensor data and operational activity reports (on a hard disk or magnetic tape) for future access and analysis by engineers and management.
- (7) The TFC shall be capable of collecting sensor data, processing it, and generating Trend Reports for future access and analysis.
- (8) The TFC shall be capable of reporting alarms (i.e., fire, power outage, equipment malfunctions, equipment exceeding specific operational ranges, etc.) both at the central point (via printer and terminal) and at each Deep Space Station (via printer and terminal).
- (9) The TFC shall be capable of sending critical alarm status reports to a secondary destination if the primary destination is not available.
- (10) The TFC shall be capable of self-diagnostics and testing to check for hardware failures and communications link breakdowns. In the event of equipment failure, the TFC will report such to the operator.
- (11) The TFC shall be capable of operating in a fail-safe configuration. This feature will allow the users to request controlled loads to be set into a specific configuration (on or off) when TFC equipment fails.

The detailed design and fabrication for the TFC at Goldstone was bid by approximately eight different commercial manufacturers. The contract was awarded to AT&T Guilford Center in Greensboro, North Carolina, which satisfied all of the requirements using a microprocessor-based commercial off-the-shelf system called Affirm III.

III. System Architecture

The TFC system consists of four basic building blocks that can be configured to match specific user needs. The four blocks are shown in Fig. 1 and are described in the following paragraphs:

- A. Central Control Unit (CCU)
- B. Local Control Unit (LCU)
- C. Sensor Control and Network Node (SCANN)
- D. Terminals and Printers

A. Central Control Unit (CCU)

The CCU, shown in Fig. 2, is an AT&T Applications Processor with an 8086 CPU, 512k of memory, and a 40 Mbyte hard disk. The current configuration of the CCU allows for communication over 12 Electronic Industries Association (EIA) standard RS232 ports and 18 Standard Serial Interface (SSI) standard RS232 ports. The EIA ports run at 1200 bps and are used to communicate over standard phone lines to the LCU's remote printers and remote terminals. The SSI ports operate at 19.2 kbps and are used as the local primary interface to the System Administrator in the form of a System Printer and Terminal and an Alarm Printer and Terminal. The CCU operates using UNIX 3.0 with the applications program (Automated Building Management [ABM]) developed by Bell Laboratories, New Jersey [1]. This package serves as the primary user interface and database management area. The ABM package is a menu driven system for editing all system databases and providing energy management features. These data-

bases include installation information, sensor descriptions, event descriptions, energy management data, and controlled load definitions. All information is centralized at the CCU for access by the System Administrator. However, it does not execute any commands or make decisions on energy management. These energy management tasks are left to the SCANNs and LCUs. After all databases are edited, they are downloaded to the LCUs and SCANNs for ABM operations. With this method of operation, the user can manage system databases and historical data without interrupting ABM operations.

B. Local Control Unit (LCU)

The LCU, as shown in Fig. 3, serves as the workhorse of the system architecture. All decision-making logic resides at the LCU and is communicated to SCANNs as required. Once database information is downloaded from the CCU it is the responsibility of the LCU to provide control and monitor information at the SCANNs under its direct command. Likewise, it is the responsibility of the LCU to report sensor status changes and diagnostic information back to the CCU at regular intervals. The LCU communicates with the CCU over two serial links: the Alarm Link and the Interactive Link. The Alarm Link is dedicated to reporting critical alarm information back to the CCU. The Interactive Link is provided for all non-critical communication with the CCU. This communication includes database downloading and diagnostic information reporting. At a relatively quiet time during ABM operation (for example, 2:00 AM) the CCU reloads the LCU with databases and the LCU in turn transmits archive data and activity reports back to the CCU for future access by the user. A second function of the LCU is to provide locations to tie binary sensors, and load controls directly into the LCU, without requiring a SCANN. Each LCU is capable of directly scanning 256 binary sensors, directly controlling 96 loads, and communicating with up to 8 SCANN units.

C. Sensor Control and Network Node (SCANN)

The SCANN, as shown in Fig. 4, is the primary front end of the system architecture. It consists of a power supply, a maintenance panel, and a card cage that can be configured for various applications. The card cage has 9 slots that can handle CPU, Communications, Binary Input, Binary Output, Analog Input, and Power Meter Input circuit cards. The SCANN can be configured as required with any combination of circuit cards (CPU and communications cards are required). These SCANNs send status messages to the Local Control Unit (LCU) whenever a change in state of a sensor is detected. If no change of state is detected, the LCU polls the SCANN for a status message at one minute intervals. These SCANNs can be located as far as 300 m from an LCU without requiring a modem or other line conditioning.

D. Terminals and Printers

The printers and terminals are the primary user interface. There are basically two types of interfaces: EIA and SSI. The EIA terminals and printers serve as remote communications ports to locations requiring a modem to communicate with the CCU. These printers and terminals serve a dual purpose: to provide remote access to the CCU and to allow direct communication with the local LCU and its associated SCANNs. In a normal configuration, these printers and terminals interface with the CCU for the user to request reports and system status. If the CCU fails or "goes down," the system can be reconfigured for the LCUs to send critical information to the local printer and terminal instead of the CCU. This is called the "Degrade Mode" of operation that serves as a suitable backup when equipment fails or communications links are broken.

Printers and terminals of the second type, SSIs, are in direct interface to the CCU. They reside within 1500 m of the CCU and can communicate over standard phone lines. The Alarm Terminal serves as the primary list device for all critical system sensors and alarm points. The Alarm Printer serves as the primary hardcopy listing device for system alarm activity. The System Administrator's Terminal is the primary location for database editing and requests for reports. The System Printer is a high speed printer that acts as the primary list device for the generation of reports requested by any user of the system.

The four TFC building blocks described above make the selected TFC system flexible, expandable, and readily configured into JPL's specific applications.

IV. System Description

The current TFC configuration is shown in Fig. 5. This configuration consists of 26 SCANNs attached to 4 LCUs communicating with one CCU. Also interfacing with the CCU are one System Administrator's Terminal, one Alarm Terminal, one System Printer, one Alarm Printer, a Remote Printer and Terminal at the Venus station (DSS 13), and a Remote Printer and Terminal at the Mars station (DSS 14). One LCU is located at each station with two located at the Mars station because of its primary activity. The SCANNs are distributed throughout the Complex at all locations where facilities data is currently required or will be required in the future. The Alarm Terminal is currently located at the Goldstone Communications Facility (GCF-10), which is the only facility that is attended 24 hours a day and 7 days a week. At this terminal, all Complex alarms are reported and acknowledged, and appropriate action taken. The Alarm Printer is located at the facility's Duty Electrician Shop to enable the resident electrician to watch the status of critical alarms. The System Printer and System Administrator's Terminals are located in the System Administrator's

office. From this location, tight control is maintained on all database editing and usage of the system. Additionally, the TFC is closely monitored for erroneous activity, hardware failures, and software difficulties. An extensive error log and activity report allows access to this information. The current configuration, therefore, is easily interfaced and readily accessible to all users who require system data.

Since the central core of TFC equipment was installed in August 1986, many different subsystems have been interfaced to the TFC. Table 1 shows the current instrumentation that interfaces to the TFC.

The primary use of the TFC is in the monitoring of critical facilities and operational equipment. At the present time, the TFC has little control activity. Most equipment at the Complex could not be controlled on an energy conservation basis due to the continuous demand for equipment usage.

V. Performance Analysis

A thorough analysis was completed on the functional performance of the TFC. This analysis consisted of a three phase collection of operational data.

The first phase of data collection consisted of developing a complete list of historical activity on the system over a period of one and one-half months between July 16, 1987, and September 1, 1987. The activity report consisted of alarm reports, analog sensor trend data, and system diagnostic activities. This data was tabulated and cross-checked with the log books from the Complex Operators. The log books recorded the day and time of the TFC activity, the action executed as a result of the activity, and the solution to the problem that caused the TFC alarm. This information is useful in evaluating the utilization of the TFC as a facilities controller device. The second level of data collection consisted of interviewing the Complex Operators, the System Administrator, and facilities personnel as to their suggestions and recommendations on the utilization of the system. This information addressed aspects of the TFC that include operability, flexibility, and design suggestions in order to improve the system utilization.

The third level of data collection consisted of inspecting a 14 month interval of system error logs. This gave an accurate measure of the reliability and availability of the system in its current configuration.

VI. Analysis Results

The first phase of evaluation, the activity report and log, concluded with the following results about the TFC:

- (1) Ninety-one TFC alarms were false due to people working on the system in alarm.
- (2) Fifty-one TFC alarms were serious alarms that required immediate response by maintenance technicians.
- (3) Eighty-two TFC alarms were "status messages" to provide useful information to facilities personnel.

As a subset of this group, there were 126 alarms due to faulty equipment generating excessive alarms. The equipment often reported error conditions more than one time.

It is noted that 41 percent of the alarms were due to people working on the equipment in alarm. This is one function of the TFC that the facilities personnel found very useful. The TFC provides a means to insure that technicians do not bring down a critical piece of equipment at an improper time. There were multiple situations during the 1.5 month evaluation where technicians were executing preventive maintenance not knowing that they were jeopardizing the proper operation of mission critical equipment. The TFC is useful in predicting these problems before they cause an operational failure.

The second phase of evaluation, personnel interviews, resulted in the following statement: The TFC is a useful facilities tool in the maintenance and troubleshooting of facilities and operational equipment. The data collection capability has been found to be useful in fine-tuning HVAC controllers. The capability of checking the status of a particular system from a remote location has been acknowledged as a useful feature. The ABM software package was found to be easy to operate with its menu driven screens. The System Administrator found that it was a simple process to add monitor instrumentation to any node of the system.

Suggested improvements to the system were provided by operations personnel as follows:

- (1) Include color graphics screen design tools to improve the operability of the monitoring capabilities.
- (2) Provide a backup power source for the SCANNs in case of commercial power outage.
- (3) Improve local diagnostic capabilities at the LCUs.
- (4) Provide screen graphing capabilities for trend reports.

The third phase of evaluation showed that the TFC, once installed and operational, does meet the availability requirements set forth in the Functional Requirements Document

(FRD 824-4).¹ The estimated availability of the TFC over the 15 month evaluation period was shown to be 99.9699 percent. The implementation phase was challenging, and the CCU, as initially received from AT&T, was unreliable. However, since the problem was traced to a faulty hard disk and the element replaced, the system has been very reliable with few noticeable hindrances.

One important feature of the TFC is its extensive remote diagnostics capabilities. All operations at the Complex can be

¹Mark IVA Technical Facilities Subsystem (1982 to 1986), JPL DSCC Subsystem Functional Requirements Document 824-4, Rev. C, Jet Propulsion Laboratory, Pasadena, California, April 1, 1984.

executed from a remote location on dial-up phone lines. This allows the Cognizant Design Engineer to assist in troubleshooting and survey activity on the system.

VII. Summary

A Technical Facilities Controller was designed and installed at the Goldstone Deep Space Communications Complex for the monitoring and controlling of facilities equipment. In the 15 month post-installation period, the TFC shows its many advantages in the operation and maintenance of facilities equipment at the Complex. As a result, it is recommended that a TFC be implemented at the overseas DSN Complexes at Canberra and Madrid.

Acknowledgments

The writers would like to acknowledge the guidance and support of the following individuals: Robert Evans, David Kuma, and Ron Casperson of the Telecommunications and Data Acquisition Program Office; Robert White, Glen Kroll, Rollin Reynolds, and Fikry Lansing of the Ground Antenna and Facilities Engineering Section; and Billy Walker and his staff of DSN Operations for their assistance throughout the design and implementation of this project.

References

- [1] *Automated Building Management (ABM) System Description*, JPL TM23910, AT&T, Issue 1, August 1983.

Table 1. Current instrumentation allocation

Subsystem	Digital sensor	Analog sensor	Load control	Power meter
Fire detection	87	0	0	0
HVAC	116	18	2	0
Water distribution	0	8	0	0
Power	12	0	2	4
Critical operational alarms	20	0	1	0
TFC self-diagnostics	32	0	2	0
Totals	267	26	7	4

ORIGINAL PAGE IS
OF POOR QUALITY

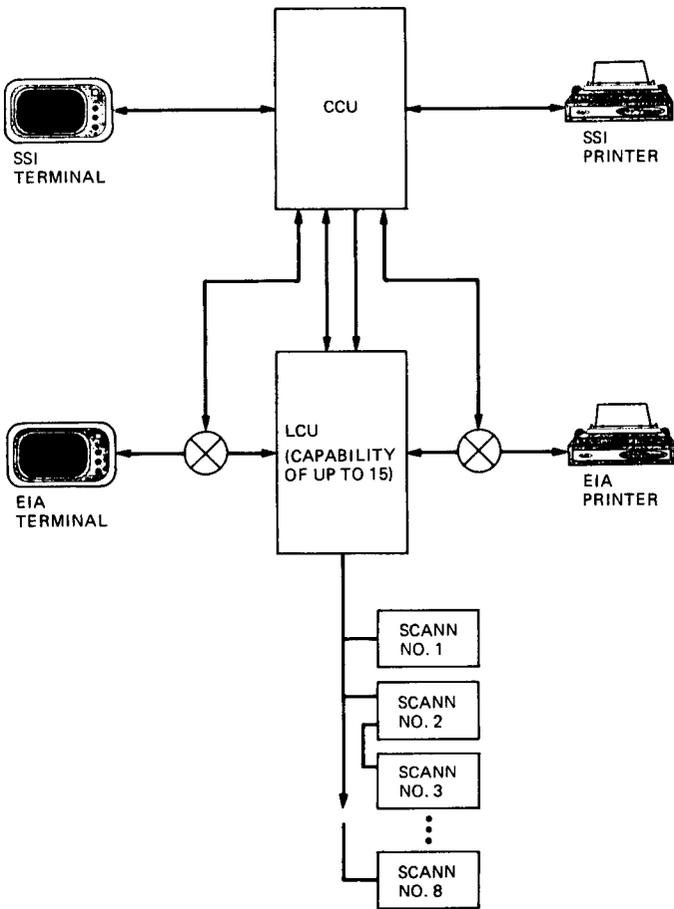


Fig. 1. TFC major elements

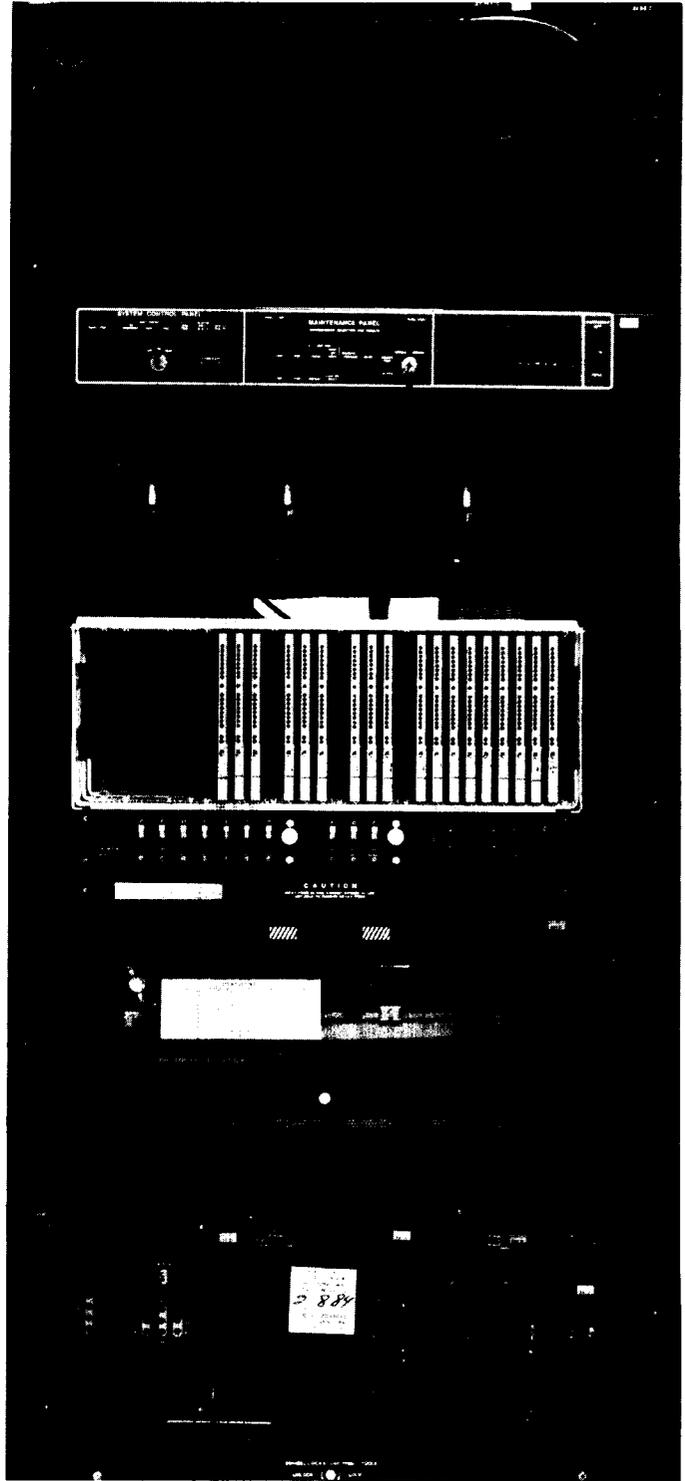


Fig. 2. Central control unit (CCU)

ORIGINAL PAGE IS
OF POOR QUALITY.

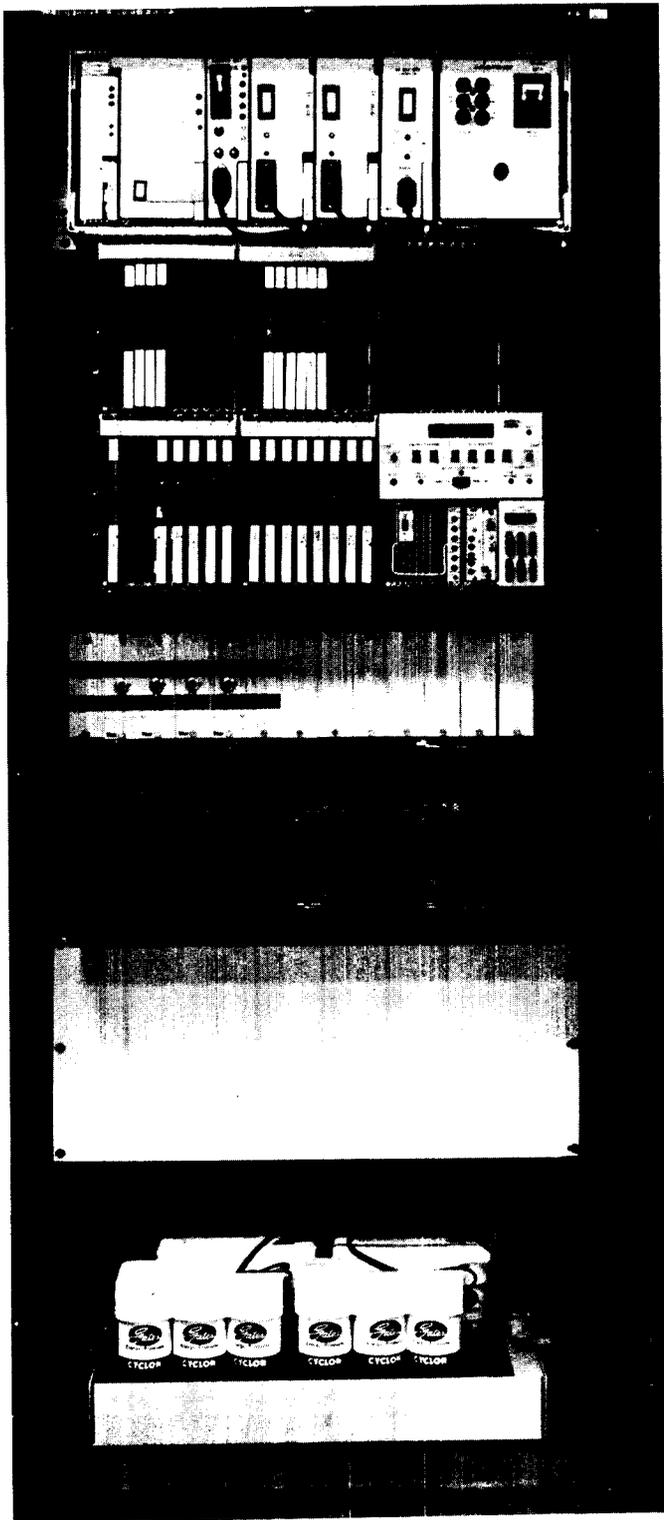


Fig. 3. Local control unit (LCU)

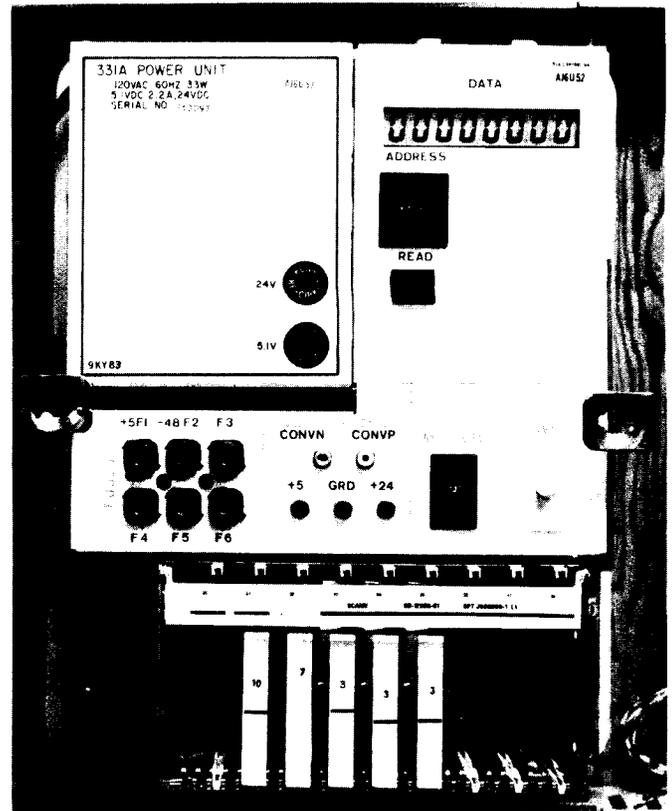


Fig. 4. Sensor control and network node (SCANN)

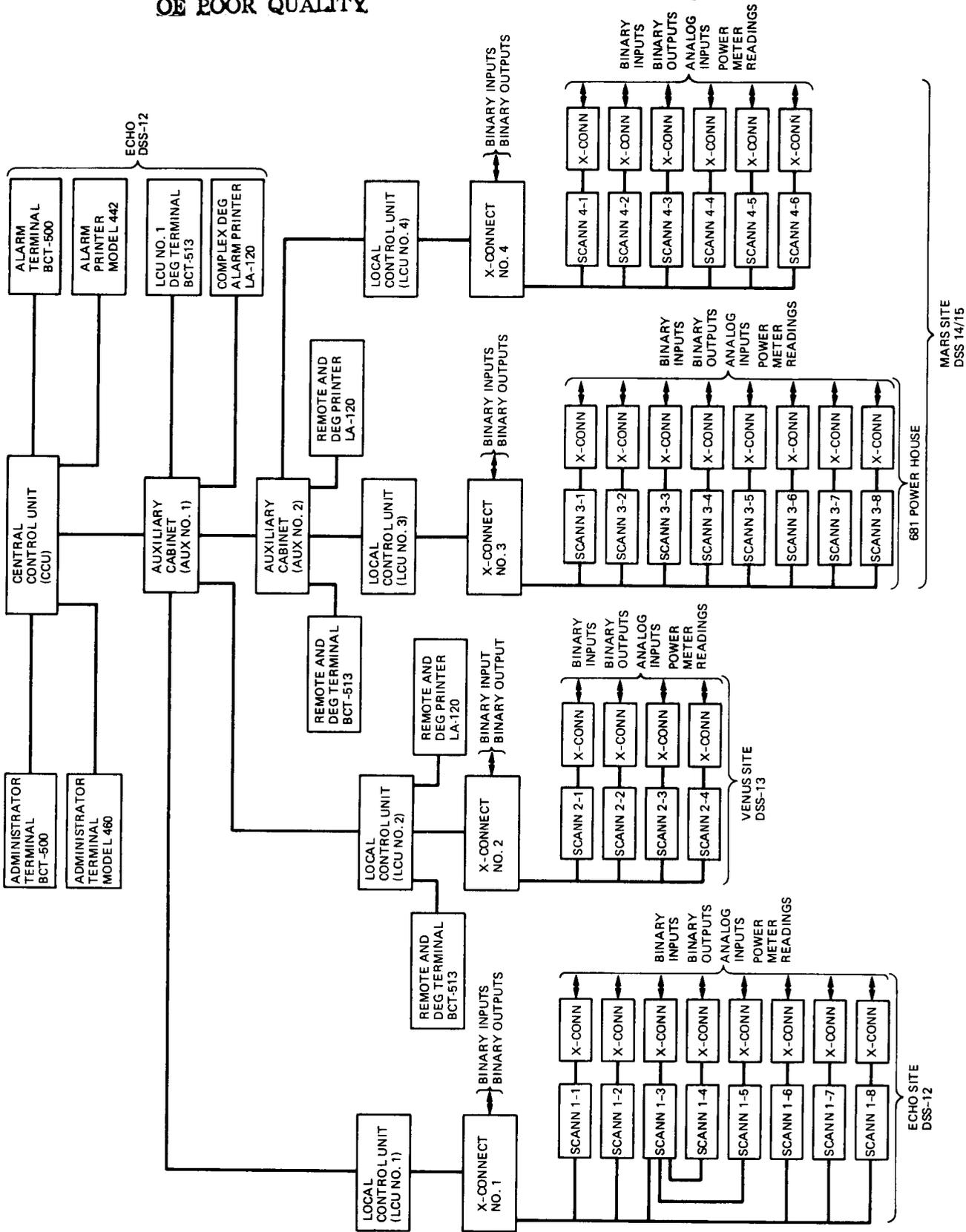


Fig. 5. Current configuration

351-33

128237

TDA Progress Report 42-92

18.

N88-18795

October-December 1987

SETI Data Controllers

R. M. Gosline

Radio Frequency and Microwave Subsystems Section

This article describes three data controllers developed for the SETI project. Two are used primarily for recording and playback of SETI data from the Radio Science Surveillance System (RSSS). The third is used as a SETI station controller for DSS 13.

I. Introduction

This article describes the SETI Data Controllers. There are three units, all based on the same hardware design. Variations in the software allow them to be used for somewhat different purposes.

II. Functional Description

The basic function of the data controllers is to record real-time data on floppy disks and to read it back. The capacity of a single disk is 512,512 bytes. In addition, the unit located in the DSS 13 control room (DR-2) must have the capability of controlling the station by sending sequences of commands through the station controller to the various subsystem controllers.

There are three data controller units, designated DR-1, DR-2, and DR-3. DR-1 and DR-3 are similar (ROM-based) units (except that DR-1 has four disk drives instead of two), but they are used in slightly different ways. DR-1 (located in the RFI van at DSS 13) is used for recording data from the Radio Science Surveillance System (RSSS). DR-3 (located at JPL) is used for playback of data and recording of sequence-of-events (SOE) files to be used at DSS 13. DR-2 (located in the DSS 13 control room) has additional capabilities for station monitoring and control.

Referring to Fig. 1, DR-1 collects real-time data from the RSSS controller and responds to data controller commands, starting with the character “#” embedded in the data stream. Note that all characters in the data stream are recorded automatically (except for command characters), provided that a write enable command has been received. DR-3 (Fig. 3) utilizes the data collection port to input command (SOE) files from the radio science VAX as data. This is because DR-3 is a ROM-based system and the command files cannot be written directly in CP/M format. DR-3 also responds to “#” commands. Real-time data is assembled in blocks of 256 bytes and written to a data disk in a sequential format to allow maximum storage capability on the floppy disk (see Table 1). Individual disk sectors contain a checksum which is transparent to the user. The data blocks may be read back a block at a time and will be preceded by a block identification number and followed by a checksum for transmission integrity.

An important feature unique to DR-1 is a power failure protection system incorporating a power fail interrupt and a non-volatile core memory. The controller will not lose data upon the occurrence of a power failure. In addition, when power is restored, the controller returns to its previous status after inserting a power failure message and count in the data stream. Should a complete restart and memory clear be desired, the operator must press the TEST button within 1 second after restoring power.

DR-2 (Fig. 2) is a RAM-disk-based system capable of supporting both the CP/M and ISIS operating systems. It has the capability of creating, reading, writing, and editing command (SOE) files to control the station through the subsystem controllers installed for the unattended station development. Higher level command capability than that available with the existing station controller is provided to allow automatic repeat of a single command at specified time intervals and looping of command sequences. In addition, a command may be preceded by a time parameter, which will cause the command to not be executed until the indicated time. In normal operation, a sequence of commands is stored in a command buffer. The buffer may be loaded from a disk file, from the DSS 13 VAX via the user port, or from the local terminal.

III. Hardware Overview

The data controllers are microprocessors assembled from standard Multibus I modules based on the 8080 CPU. They are housed in individual seven-slot chassis with internal power supplies. Each controller has a dual floppy disk drive using Shugart 801-R drives, housed in a separate chassis, and operating in the double density mode.

DR-1 and DR-3 are ROM-based systems using the Intel SBC 80-20/4 single board computer. The program is contained on board in 2716 PROMs. A Zendex ZX-200A disk controller is used to control the disk drives. A 16-kbyte RAM card is provided to allow the disk controller to operate in DMA mode. The 80-20/4 contains a serial port (designated channel A) that is used for sending and receiving data. A serial port module has been added to DR-1 to allow use of a monitor terminal, modem, and printer.

DR-2 contains 64 kbytes of RAM and is a RAM disk-based system. The program must be loaded from disk after booting the CP/M operating system. The National BLC 80/204 single board computer is used as the CPU. This card allows more flexibility in the use of an on-board monitor ROM whose contents can be moved to high RAM and then shadowed by a software command to allow use of standard CP/M or ISIS operating systems. A BLC-534 communications expansion module provides four additional serial ports.

IV. Software Overview

The software for all controllers is written in the PL/M-80 language, assembled, linked, and located under the ISIS operating system. The load modules are copied to a CP/M disk which can be loaded under the CP/M operating system for DR-2 or burned in PROM for DR-1 and DR-3. Other utilities such as editors can be loaded to facilitate modification or

creation of command files. The software for all three controllers was developed in DR-2. The principles of top-down structured programming were followed.

V. Command Descriptions

Except for loops, all commands are processed in sequence, i.e., they are not sorted or reordered. Table 2 is a list of allowable commands from the various ports. R/T commands are processed immediately. Command buffer commands are processed immediately if not preceded by a time parameter. If preceded by a time parameter, the command is processed as soon as the following condition exists:

$$\text{Station Time} \geq \text{Command Time}$$

Note that a timed command or a delay command may hold up the execution of the next command, even if it is in the immediate mode.

A. Subsystem Commands

The subsystem commands are explained in the "Unattended Station Controller Operator's Handbook."

B. Data Controller "#" Commands

Data Controller commands all contain six bytes and have the form #xxx"CR/LF."

#IN0	Format disk in drive 0 as a data disk. Set current drive to 0.
#IN1	Format disk in drive 1 as a data disk. Set current drive to 1.
#IN2	Format disk in drive 2 as a data disk. Set current drive to 2.
#IN3	Format disk in drive 3 as a data disk. Set current drive to 3.
#INA	Format disks in all four drives (DR-1 only), starting with drive 0 and continuing until encountering an empty drive.
RDA	Read entire data file to printer port, starting with drive 0 and switching drives until no disk is found or a hex E5 character is found. The output does not include a block ID or checksum.
#BL0	Set the block pointer for drive 0 to the value specified in the next four ASCII characters. The value must be in the range 0001-2002 and must contain exactly four characters. Set the current drive to 0.
#BL1	Set the block pointer and current drive for drive 1.

#BL2 Set the block pointer and current drive for drive 2.

#BL3 Set the block pointer and current drive for drive 3.

#RDO Set the current drive to 0. Read block from current drive and increment block pointer by 1. The pointer wraps around from 2002 to 0001. If the pointer wraps around, the current drive switches to 1.

#RD1 Set the current drive to 1. Read block from current drive and increment block pointer. If the pointer wraps around, the current drive switches to 2.

#RD2 Set the current drive to 2. Read block from current drive and increment block pointer. If the pointer wraps around, the current drive switches to 3.

#RD3 Set the current drive to 3. Read block from current drive and increment block pointer. If the pointer wraps around, the current drive switches to 0.

#EOF Fill any space in the current data block with hex E5 characters and write it to the current drive. Increment the block pointer as done in the #RDO and #RD1 commands.

#RST Reset the data controller program and memory to the initial condition after boot or load.

#STR Return data controller status.

#WE0 Enable drive 0 to allow write operations.

#WE1 Enable drive 1 to allow write operations.

#WE2 Enable drive 2 to allow write operations.

#WE3 Enable drive 3 to allow write operations.

#WD0 Disable drive 0 write and format operations.

#WD1 Disable drive 1 write and format operations.

#WD2 Disable drive 2 write and format operations.

#WD3 Disable drive 3 write and format operations.

#PTA Print all data in RDA format.

C. DR-2 Control Commands

START Begin processing command buffer (run mode).

STOP Stop processing command buffer.

DISPLAY [n] Display 24 lines from line *n* of command buffer. Move command line pointer to

line *n*. If *n* is omitted, it displays from current command line.

DELETE [n] Delete *n* lines in command buffer from the current pointer. If *n* is omitted, one line is deleted.

INSERT Insert lines before cursor until ESC character.

"ESC" Terminate INSERT.

LOOP *n* Repeat from here to ENDLOOP *n* times. Nesting is not permitted. May not follow an RPT*n* command.

ENDLOOP Signifies end of loop.

DELAY *n* Wait *n* seconds before processing next command.

HELP [n] Help menus.

LOAD file Load command file from CP/M disk. Prompt for correct disk. The maximum file size is 32 kbytes. Prompt for correct disk type.

SAVE file Write command file to disk. Not permitted in run mode. Prompt for correct disk type.

RPT*n* *S* Repeat next command at interval *S* (seconds). The maximum allowable is 65535. The command may be canceled with RPT*n* 0 (*n* = 0, 1, . . . , 9). Nesting is not allowed for the same *n*.

TOFF *T* Add a positive time offset *T* to all command times and data times as they are processed until stop command. *T* is entered in day/hour/minute/second format or the character @ for current station time. The day result is not corrected for a value in excess of a year. The day total and carry may exceed 365. The command in the buffer is not changed, but the disk data and terminal display include the offset.

DATON Start writing data to disk.

DATOFF Stop writing data to disk.

STAIUS Report the current command buffer status.

CPM2DAT file Copy a CP/M file from drive 0 to an empty data disk on drive 1, then execute an automatic #EOF.

DAT2CPM Copy a data disk on drive 0 to CP/M disk on drive 1. The file will be broken up into

32,768 byte CP/M files or at the first hex E5 character. The file name will be taken from characters 2 through 9 of the first data block, with the file extension assigned numerically starting with 001. If the CP/M disk does not contain enough room

Y
N

for the next file, a prompt is issued for another disk on drive 1.

Affirmative response to a query.

Negative response to a query.

Table 1. Data block format

Byte	Contents
1-4	Block identification number—must be in the range 0001-2002 and contain only ASCII numerals
5-260	256 bytes of ASCII data containing only printable ASCII characters
261-262	Checksum formatted as 2 hex ASCII bytes (high-order first)*

*The checksum is computed by performing an "exclusive OR" operation over the first 260 bytes of the block. The resultant byte is split into two hex ASCII digits. For example, a checksum of 91 (hex 5B) results in the ASCII characters "5" and "B" for characters 261 and 262, respectively.

Table 2. Allowable commands

Command	Channel A	Channel E	Channel F	Command buffer
#INn	X	X	X	X
#INA	X	X	X	X
#RDA	X	X	X	X
#BLn	X	X	X	X
#RDn	X	X	X	X
#EOF	X	X	X	X
#RST	X	X	X	X
#STR	X	X	X	X
#WEn	X	X	X	X
#WDn	X	X	X	X
#PTA	X	X	X	
START		X	X	
STOP		X	X	X
DISPLAY [N]		X	X	
DELETE [N]		X	X	
INSERT		X	X	
"ESC" (Hex 1B)		X	X	
LOOP N				X
ENDLOOP				X
DELAY N				X
HELP [N]		X	X	
LOAD file		X	X	X
SAVE file		X	X	X
RPT[N] S				X
TOFF T		X	X	X
DATON		X	X	X
DATOFF		X	X	X
STATUS		X	X	X
CPM2DAT file		X		
DAT2CPM		X		
UON		X		
UOFF		X	X	X
Y		X	X	X
N		X	X	X

n = 0, 1, 2, 3.

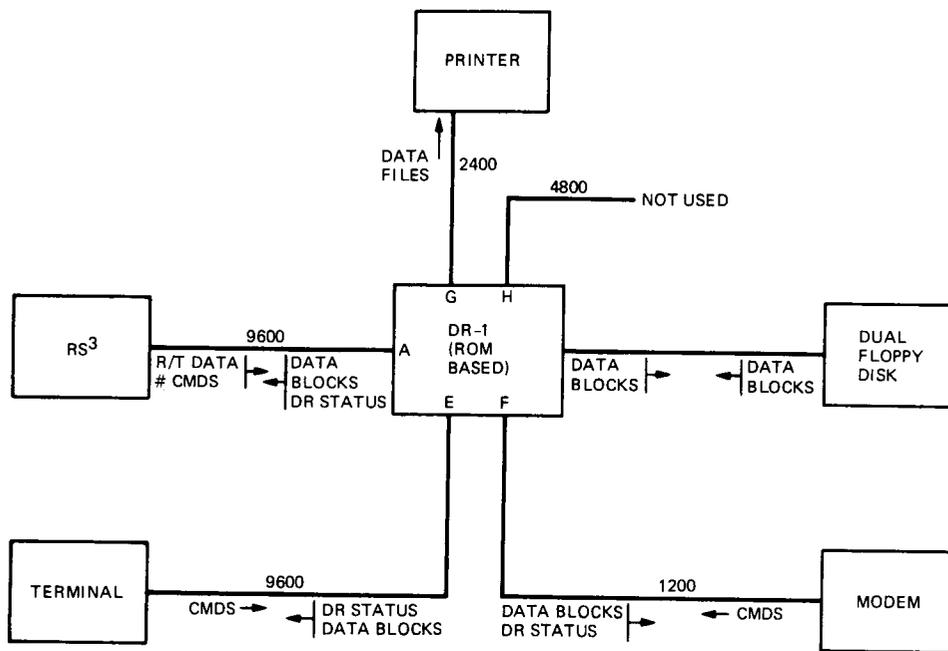
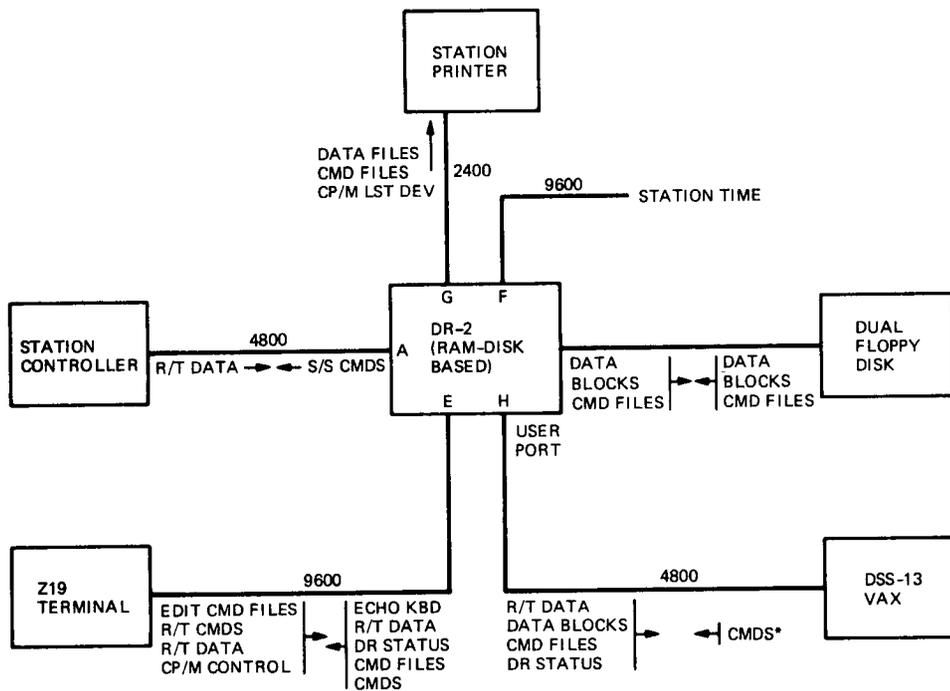


Fig. 1. Functional block diagram of DR-1



*ALLOWS TRANSFER OF COMMAND FILES WITH INSERT COMMAND

Fig. 2. Functional block diagram of DR-2

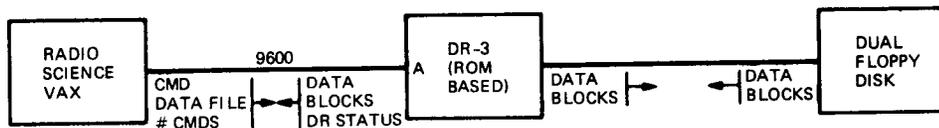


Fig. 3. Functional block diagram of DR-3

N88-18796

322-32 ✓

128258

76

Simulated Performance of an Order Statistic Threshold Strategy for Detection of Narrowband Signals

E. Satorius

Communications Systems Research Section

R. Brady and W. Deich

Image Processing Applications and Development Section

S. Gulkis and E. Olsen

Space Physics and Astrophysics Section

The application of order statistics to signal detection is becoming an increasingly active area of research. This is due to the inherent robustness of rank estimators in the presence of large outliers that would significantly degrade more conventional mean-level-based detection systems. In this article, a detection strategy is presented in which the threshold estimate is obtained using order statistics. The performance of this algorithm in the presence of simulated interference and broadband noise is evaluated. In this way, the robustness of the proposed strategy in the presence of the interference can be fully assessed as a function of the interference, noise, and detector parameters.

I. Introduction

Development of a two million channel, FFT-based narrowband detection processor is currently under way at JPL for use in various applications of the Deep Space Network [1], [2]. It will also serve as a prototype for the Search for Extraterrestrial Intelligence (SETI) Sky Survey Processor [3]. The system is being designed to process contiguous spectra at a throughput rate of 40 MHz. The system output consists of detected spectral intervals. Each interval is composed of a run of one or more contiguous spectral bins for which the asso-

ciated power levels all exceed the system threshold. The average power level, width, and location of each detected spectral interval are computed and passed along to the system computer which performs the final signal and interference assessment.

The key parameter in this system is the threshold level. For effective system performance, it is desirable that the threshold be adaptive to accommodate a typically time varying background (thermal) noise level, and that it be as insensitive as

possible to the presence of both narrowband signal and interference components that lie within the threshold estimation window. Finally, it is desired that the threshold level be stable, i.e., that it exhibit a small variance. Toward these ends, we have considered an order-statistic-based threshold estimation scheme wherein the system threshold is a constant times a linear combination of successive n th order statistics computed from the power spectral data out of the FFT (see Section II). The scaling constant controls the false alarm rate.

The analysis and application of order statistics in general has become an important area of research [4]–[6], and the specific application of order statistics to signal detection is currently receiving some attention [7]. This is due to the inherent robustness of order statistics in the presence of large outliers (e.g., narrowband interference) that would significantly degrade the sensitivity of more conventional mean-level-based detection systems. The real issue in the application of order statistics to signal detection is the system performance for a given interference environment. An analysis of an order-statistic-based detection system presented in [7] clearly demonstrates the robustness of order statistic threshold estimators to the presence of a single narrowband interferer within the estimation window. Of course, in practice there will typically be multiple interferers with different amplitudes and bandwidths within the window depending on the specific interference environment.

This article summarizes the results of a preliminary computer-aided simulation analysis that has been carried out to evaluate system performance in the presence of interference. In performing this analysis, the interference environment has been simulated based on the results of limited survey data (collected between 1 and 2 GHz) which provides the percentage of spectral bins contaminated by RF interference (RFI) as a function of the RFI power level. In addition to RFI, a broadband Gaussian system noise component is included in this analysis. Ideally, the system threshold level will reflect the spectral level of the broadband system noise and not the RFI.

This system performance evaluation is carried out as a function of the system noise level relative to the RFI. The results of this analysis not only serve to assess system performance as a function of various system parameters, but also provide guidelines for choosing various system design parameters to enhance system performance.

II. Detector and RFI Models

The basic detector system model currently under consideration is depicted in Fig. 1. Here the digitized input data are transformed into the frequency domain via an FFT processor and the power in each FFT bin is accumulated over a specified

number of transforms. The resulting accumulated power data is then split into two paths. The direct-through path is fed into a 5-point convolutional filter which forms the convolution of successive accumulated spectra with a 5-point finite impulse response (FIR) filter. Specifically, let $X_i(k)$ denote the level of the i th successive accumulated power spectrum at the k th spectral (FFT) bin. Then the output from the 5-point convolutional filter is given by:

$$Y_i(k) = \sum_{j=0}^4 w_j X_{i-j}(k)$$

where the w_j are the FIR filter weights. As discussed in [3], the filter weights depend on the characteristics of the receive antenna beam pattern and are matched to the expected signature of a fixed source in the sky as it is traversed by the receive beam. Furthermore, a 5-coefficient FIR filter turns out to be sufficient to minimize signal-to-noise ratio (SNR) losses induced by the antenna beam in conjunction with the finite time interval between successive accumulated spectra [3]. For purposes of this analysis, the FIR filter weights are considered to be a set of fixed constants downloaded from the system computer.

In addition to being directly convolved with the FIR filter, the accumulated power spectral data are also utilized in determining the system threshold level. As indicated in Fig. 1, the threshold determination is composed of three steps: (1) compute the n th smallest power level; (2) convolve successive n th smallest power levels with the FIR filter; and (3) multiply the result from (2) by a fixed gain constant. The result of (2) is to further smooth the n th order statistics computed in (1) in a manner which is perfectly consistent with the convolution of the power accumulation data in the direct-through path. Note that this smoothing operation also reduces the variance of the order statistics. The final step effectively determines the number of detected noise intervals (or false alarm rate) out of the detector system. The gain constant in (3) is typically chosen based on a system-noise-only (no RFI) assumption. The goal of this analysis is to determine the threshold stability as well as the number of spectral interval detections in the presence of RFI.

After convolution and threshold level determination, the convolved data are thresholded and information concerning detected spectral intervals (interval width, location, and average power level) is passed on to the system computer. Ideally, the thresholding serves to discard most of the noise data (so that the system computer is not overloaded) while simultaneously retaining the desired signal information. The purpose of passing along detected intervals and not individual spectral bin detections is to reduce the amount of hit data which will arise from broadband interference sources with

bandwidths well in excess of the FFT bin resolution (≈ 30 Hz). Interference-related hit data which is passed on to the computer can then be identified (e.g., based on frequency or time discrimination) and eliminated from further analysis.

The RFI model developed for the system performance evaluation is based on limited survey data collected in the 1–2 GHz band. In collecting this data, a spectral resolution width of 10 kHz was used, and the average number of threshold detections attributed to RFI sources over a 1 GHz bandwidth was computed as a function of the threshold level. The fraction of spectral bins contaminated by RFI fit a power law model as the threshold level decreased over the range from -80 dBm to -120 dBm. Figure 2 represents the least squares fit to this data. In Fig. 2, the log (base 10) of the fraction of RFI-contaminated bins is plotted versus the RFI threshold level. For levels above -80 dBm, the fraction of spectral bins contaminated by the RFI is assumed constant. Below -120 dBm, the survey data has been extrapolated exponentially to -140 dBm. Below -140 dBm a constant profile of approximately 15 percent RFI contamination is assumed. The simulation results presented in Section III are based on this RFI density profile.

There are two basic limitations associated with this model. First, there are currently no available RFI data measurements below approximately -120 dBm. Even for a 10 kHz spectral resolution, this is well above the thermal noise level (≈ -150 dBm assuming a nominal 10 K system temperature and 10 kHz bandwidth). It was thus necessary to extrapolate the data into the low noise regions of interest as indicated in Fig. 2. A second limitation is that the RFI data used to construct this model profile have been collected using a spectral resolution (10 kHz) which is far coarser than the system goal (≈ 30 Hz). Consequently, the RFI realizations based on this density profile will differ significantly from those corresponding to a narrowband RFI profile. Nevertheless, system performance results based on this model do highlight the RFI model attributes which most critically impact system performance in general.

III. System Performance Assessment

In the evaluation of system performance, a computer simulation test bed has been developed which generates realizations of RFI based on the amplitude density model depicted in Fig. 2. Specifically, the RFI amplitude range has been quantized into 2.5 dBm intervals, and the appropriate number of RFI components has been injected into each interval. Furthermore, the RFI components have been distributed into non-overlapping frequency bin intervals with all of the RFI com-

ponents in a frequency interval having amplitudes lying within a given 2.5 dBm amplitude interval. Each RFI frequency interval is randomly positioned across the total number of spectral bins selected (i.e., total instantaneous bandwidth), and the width of each RFI interval is chosen randomly up to a maximum of 40 contiguous spectral bins. The resulting distribution of the RFI in frequency is termed the RFI "mask." A sample RFI mask is presented in Fig. 3, where the average interference-to-noise ratio (INR) is plotted across a total of 4096 spectral bins corresponding to a system noise level of -110 dBm. The power law increase in the number of RFI components with decreasing INR is clearly observed.

Each simulation run is composed of multiple realizations of RFI and additive broadband system noise corresponding to a fixed RFI mask (one independent mask per simulation run). The RFI amplitude (dBm) is uniformly randomized within each 2.5 dBm amplitude interval, and the phases of all RFI components are uniformly randomized over $[0, 2\pi)$ once every spectral accumulation cycle. Independent broadband noise realizations (generated in the frequency domain) are computed for each power spectrum input to the accumulator.

The simulation input parameters include (1) total number of spectral bins (nominally 4096); (2) number of spectra per accumulation cycle (nominally 8); (3) average system noise level within a spectral bin (this varied between -100 and -150 dBm); (4) rank number, n , for the order statistic ($n = 10$ or 60); (5) gain constant for computing the system threshold corresponding to a 0.1 percent false alarm rate in the absence of RFI; and (6) total number of accumulated spectra (nominally 1000 per run). All of these parameters, including the RFI mask, are held constant during a simulation run. In addition, the 5 convolutional filter weights are always fixed (at the values 0.64, 0.89, 1.0, 0.89, and 0.64).

The two primary outputs from the system simulations are the n th order statistic and the number of detected noise intervals, averaged over all realizations, as a function of the system noise level. Plots of the order statistics (normalized by their respective means in the absence of RFI) are presented in Fig. 4 corresponding to the nominal simulation input parameters given above. As is seen, when the noise level is well above the majority of RFI components, i.e., above -90 dBm, then the spectral bins are dominated by system noise and the resulting n th order statistics approach those for the noise-only distribution. Conversely, as the system noise level decreases to well below the smallest RFI component, i.e., below -140 dBm, then the RFI components contaminate a fixed number of spectral bins and the order statistics increase due to the corresponding reduction in the number of noise-only bins. This is clearly observed in Fig. 4, where a significant inflation of both

the 10th and the 60th smallest order statistics occurs as the noise level decreases. The net result of this inflation is a loss in detector sensitivity. However, the sensitivity loss associated with either order statistic is much less than would occur using a conventional average power estimator [7].

Plots of the average number of detected spectral intervals corresponding to both the 10th smallest and 60th smallest order statistics are presented in Fig. 5. As is seen for both cases, the average number of detected intervals increases from the noise-only limit of approximately 4 (0.1 percent of 4096 spectral bins) to over 80 at the -140 dBm noise level and then back to approximately 50 for noise levels below -150 dBm. At these lower levels, the detected spectral intervals are composed almost exclusively of the RFI frequency intervals—an average of 50 such intervals were generated in the RFI masks used for these runs. The number of additional detected noise-only intervals in this case is limited by the inflation of the n th order statistics as noted above. As the system noise level increases to -140 dBm, a level which corresponds to the majority of the RFI components (see Fig. 2), it interferes with the RFI to produce random “splittings” of the low-level, RFI frequency intervals. These splittings are manifested as an increase in the number of detected noise intervals to over 20 times that expected in the absence of RFI. Such an increase could impact the ability of the system computer to process all of the detected hit data. A more precise assessment of the impact of RFI on system performance will, in turn, require a more complete set of RFI survey data, which is clearly an important area for future investigation.

IV. Conclusions

Although the results of this preliminary simulation analysis depend critically on the assumed RFI density model, some general conclusions can be made. In particular, it is noted that system performance depends most critically on the distribution of RFI components at levels comparable to or greater than the broadband system noise level. RFI components well below the system noise level do not significantly impact system performance. RFI components well in excess of the system noise level contaminate a fixed number of spectral bins and thus produce a significant inflation of the system threshold due to the decrease in the number of noise-only spectral bins. This inflation has the effect of lowering the average number of detected noise-only spectral intervals as well as reducing detector sensitivity. Large RFI components also increase the total number of detected spectral intervals depending on the number of RFI intervals present. Note that the number of RFI intervals will, in turn, depend on the frequency distribution of the RFI. A small number of strong broadband RFI sources will not significantly increase the total number of detected spectral intervals, whereas a large number of strong narrowband RFI sources will produce a significant increase in the number of interval detections. Furthermore, as the RFI level approaches the system noise, additional spectral interval detections will occur due to RFI interval splitting caused by the interaction of the broadband system noise with the RFI. This splitting phenomenon will occur regardless of whether the RFI is narrow or broadband. Further analysis is currently being carried out to assess system performance in more realistic RFI environments.

References

- [1] M. Quirk, H. Wilck, and M. Grimm, "A Wide-Band, High-Resolution Spectrum Analyzer," *TDA Progress Report 42-83*, vol. July-September 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 180-190, November 15, 1985.
- [2] M. Garyantes, M. Grimm, H. Wilck, and E. Satorius, "Two Megachannel High Resolution Digital Spectrum Analyzer for SETI Applications," presented at the National Radio Science Meeting, Boulder, Colorado, January 12-15, 1987.
- [3] J. Solomon, W. Lawton, M. Quirk, and E. Olsen, "Signal Detection Strategy for the SETI All Sky Survey," *TDA Progress Report 42-83*, vol. July-September 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 191-208, November 15, 1985.
- [4] N. Gallagher, Jr., and G. Wise, "A Theoretical Analysis of the Properties of Median Filters," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-29, pp. 1136-1141, December 1981.
- [5] B. Justusson, "Median Filtering Statistical Properties," in *Two-Dimensional Digital Signal Processing II*, Chapter 8, pp. 161-196, T. Huang (ed.), New York: Springer-Verlag, 1981.
- [6] A. Bovik, T. Huang, and D. Munson, Jr., "A Generalization of Median Filtering Using Linear Combinations of Order Statistics," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-31, pp. 1342-1350, December 1983.
- [7] K. Wong and C. Shuang, "Detection of Narrow-Band Sonar Signals Using Order Statistical Filters," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-35, pp. 597-613, May 1987.

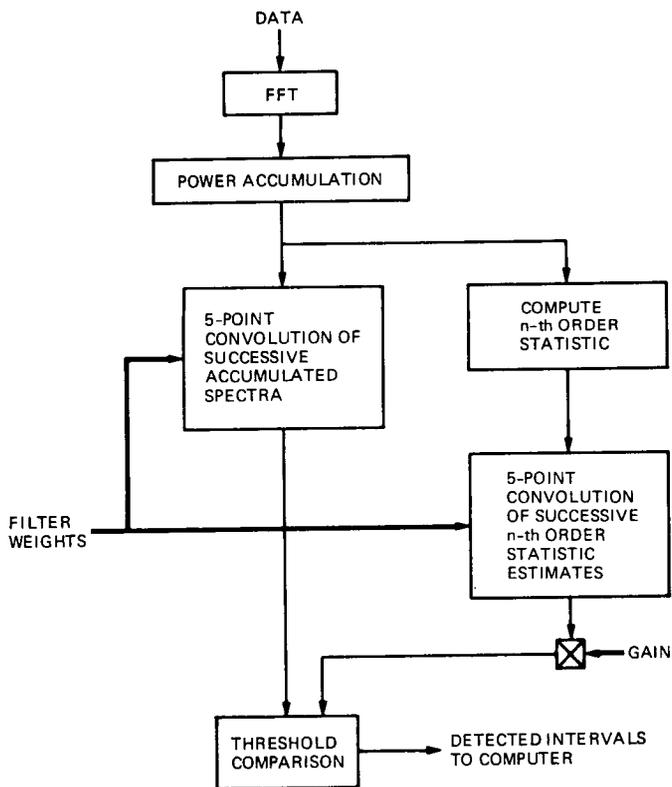


Fig. 1. Detector system model (heavy lines denote paths for detector parameters which are downloaded from the system computer)

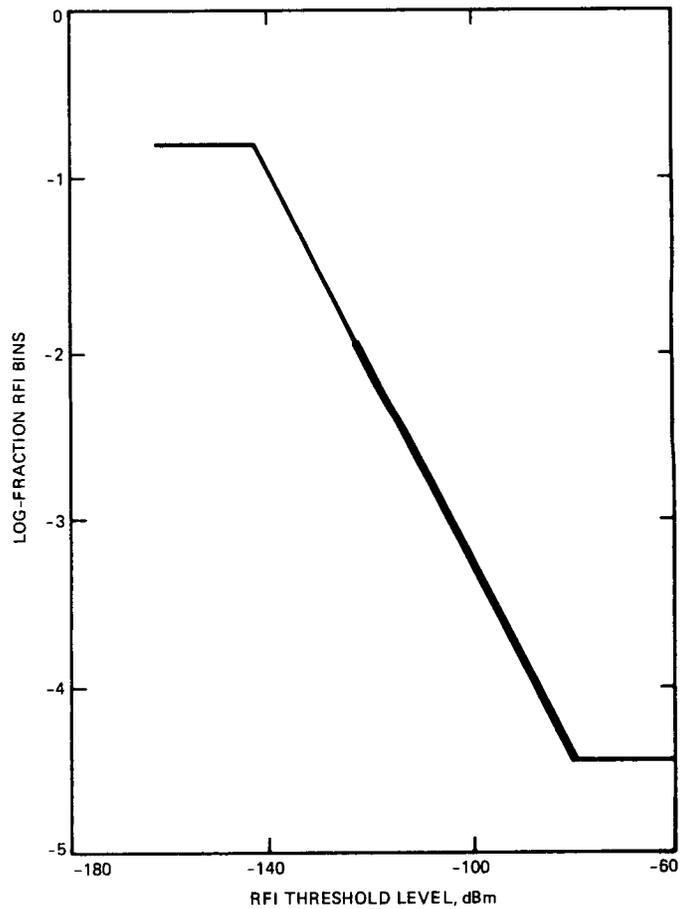


Fig. 2. RFI model density profile (heavy solid line denotes the region of valid survey data)

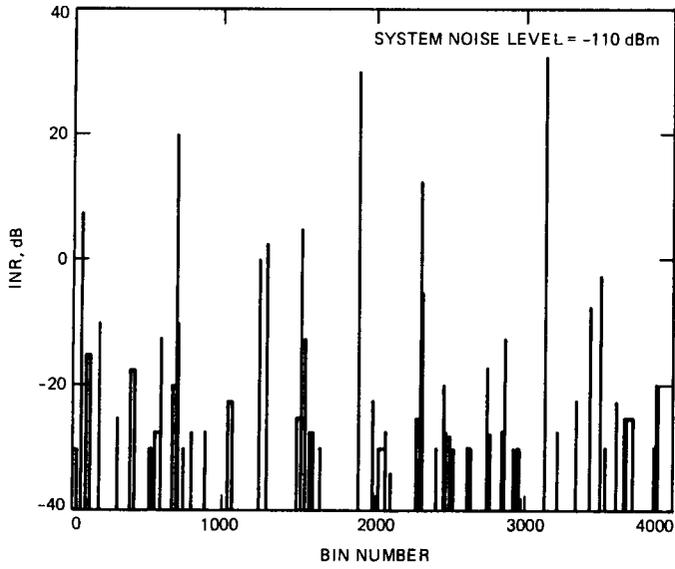


Fig. 3. Sample RFI mask

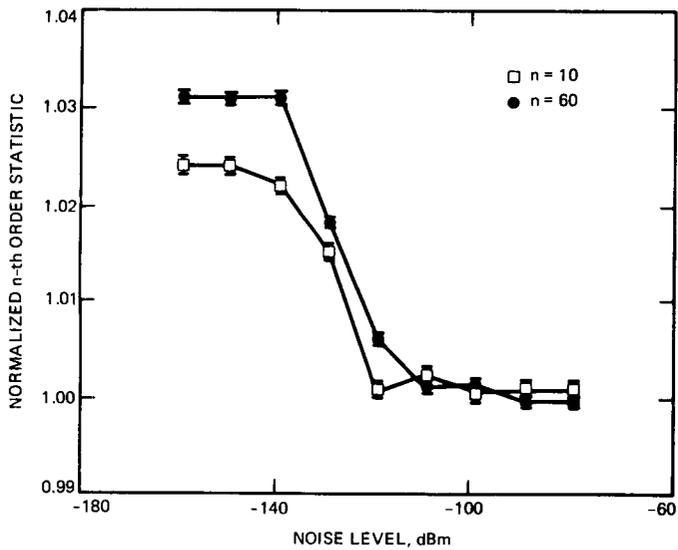


Fig. 4. Average n th order statistics (normalized by expected mean values in the absence of RFI)

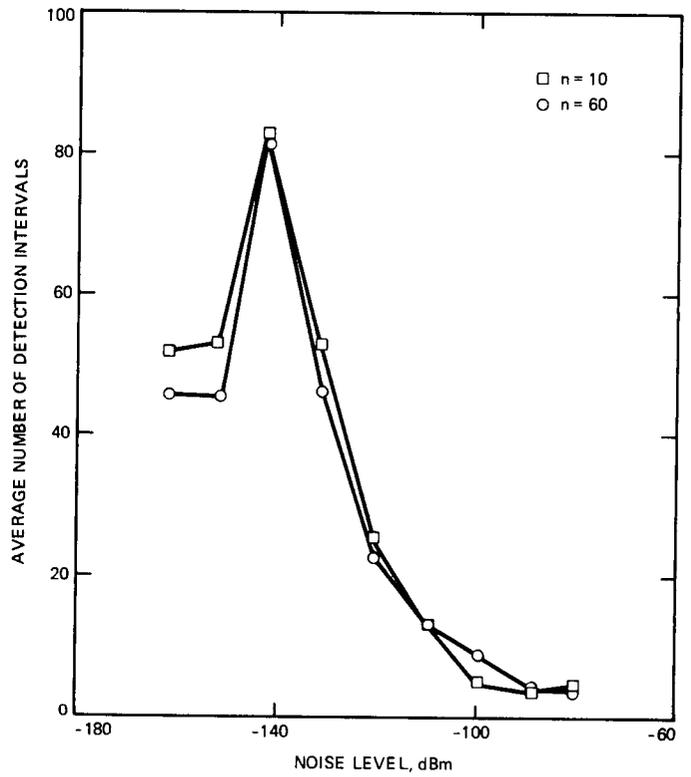


Fig. 5. Average number of detected spectral intervals

omit to END

Author Index,¹ 1987

The Telecommunications and Data Acquisition Progress Report

42-89, January-March, May 15, 1987

42-90, April-June, August 15, 1987

42-91, July-September, November 15, 1987

42-92, October-December, February 15, 1988

Abdel-Ghaffar, K.

- 42-90 Construction for Finite-State Codes, pp. 42-49.
See Pollara, F.

Aguilera, C. S. R.

- 42-92 Frame Synchronization Performance and Analysis,
pp. 110-116.
C. S. R. Aguilera, L. Swanson, and G. H. Pitt III

Annis, J.

- 42-91 The Atmosphere of Mars and Optical Communica-
tions, pp. 124-132.

Belk, P.

- 42-91 A Procedural Method for the Efficient Implementa-
tion of Full-Custom VLSI Designs, pp. 235-246.
P. Belk and N. Hickey

Bhanji, A.

- 42-90 Reduction of Ground Noise in the Transmitter
Crowbar Instrumentation System by the Use
of Baluns and Other Noise Rejection Methods,
pp. 27-31.

See Daeges, J.

- 42-91 Inductance Effects in the High-Power Transmitter
Crowbar System, pp. 81-88.

See Daeges, J.

Border, J. S.

- 42-91 A Demonstration of High Precision GPS Orbit
Determination for Geodetic Applications, pp. 1-22.

See Lichten, S. M.

Brady, R.

- 42-92 Simulated Performance of an Order Statistic Thresh-
old Strategy for Detection of Narrowband Signals,
pp. 191-197.

See Satorius, E.

Britcliffe, M.

- 42-90 Sorption Compressor/Mechanical Expander Hybrid
Refrigeration, pp. 21-26.

See Jones, J. A.

- 42-91 Traveling-Wave Maser Closed-Cycle Refrigerator
Data Acquisition and Display System, pp. 304-311.

See Fowler, L.

- 42-91 Two-Watt, 4-Kelvin Closed Cycle Refrigerator
Performance, pp. 312-317.

Brokl, S. S.

- 42-90 Survey of Current Correlators and Applications,
pp. 54-62.

See Lahmeyer, C. R.

Brown, D. H.

- 42-89 DSN Advanced Receiver: Breadboard Description
and Test Results, pp. 48-66.

D. H. Brown and W. J. Hurd

¹In case of joint authorship, the reader is referred to the citation under the first author where all authors of the article are listed.

42-91 Optimized Tracking of RF Carriers With Phase Noise, Including Pioneer 10 Results, pp. 141-157.
See Vilnrotter, V. A.

Brundage, W. D.

42-92 X-Band System Performance of the Very Large Array, pp. 123-137.
See Ulvestad, J. S.

Brunzie, T. J.

42-92 A Noise-Adding Radiometer for the Parkes Antenna, pp. 117-122.

Cardenas, R.

42-91 The 8.4-GHz Low-Noise Maser Pump Source Assembly, pp. 278-284.

Cha, A. G.

42-89 Z-Corrections for DSN 70-Meter Antenna Ranging Calibration, pp. 77-82.

Cheung, K. M.

42-91 More on the Decoder Error Probability for Reed-Solomon Codes, pp. 213-223.

42-92 A Labeling Procedure for Linear Finite-State Codes, pp. 50-55.

42-92 A Lower Bound for the Decoder Error Probability of the Linear MDS Code, pp. 43-49.

42-92 Further Results on Finite-State Codes, pp. 56-62.
See Pollara, F.

Chuang, K. L.

42-92 An Optimal Structure for a 34-Meter Millimeter-Wave Center-Fed BWG Antenna: The "Cross-Box" Concept, pp. 27-39.

Clements, P. A.

42-89 Results of Using the Global Positioning System to Maintain the Time and Frequency Synchronization in the Deep Space Network, pp. 67-72.

P. A. Clements, A. Kirk, and R. Unglaub

Daeges, J.

42-90 Reduction of Ground Noise in the Transmitter Crowbar Instrumentation System by the Use of Baluns and Other Noise Rejection Methods, pp. 27-31.

J. Daeges and A. Bhanji

42-91 Inductance Effects in the High-Power Transmitter Crowbar System, pp. 81-88.

J. Daeges and A. Bhanji

Daggett, D. L.

41-91 A Cold Ejector for Closed-Cycle Helium Refrigerators, pp. 102-111.

See Johnson, D. L.

Deich, W.

42-92 Simulated Performance of an Order Statistic Threshold Strategy for Detection of Narrowband Signals, pp. 191-197.

See Satorius, E.

Deutsch, L. J.

42-89 A VLSI Single Chip (255, 223) Reed-Solomon Encoder With Interleaver, pp. 41-47.

See Hsu, I.-S.

42-90 A Comparison of VLSI Architecture of Finite Field Multipliers Using Dual, Normal, or Standard Basis, pp. 63-75.

See Hsu, I.-S.

42-91 The Design Plan of a VLSI Single Chip (255, 223) Reed-Solomon Decoder, pp. 186-199.

See Hsu, I.-S.

42-91 On the VLSI Design of a Pipeline Reed-Solomon Decoder Using Systolic Arrays, pp. 224-234.

See Shao, H. M.

42-92 A Comparison of VLSI Architectures for Time and Transform Domain Decoding of Reed-Solomon Codes, pp. 63-81.

See Hsu, I.-S.

Dick, G. J.

42-91 Atomic Frequency Standards for Ultra-High-Frequency Stability, pp. 67-72.

See Maleki, L.

42-91 Measurement and Analysis of Cryogenic Sapphire Dielectric Resonators and DROs, pp. 73-80.

42-91 Fast Autotuning of a Hydrogen Maser by Cavity Q Modulation, pp. 295-303.

G. J. Dick and T. K. Tucker

42-92 The JPL Trapped Mercury Ion Frequency Standard, pp. 13-19.

See Prestage, J. D.

Eastman, W. L.

42-91 A Simplified Procedure for Correcting Both Errors and Erasures of a Reed-Solomon Code Using the Euclidean Algorithm, pp. 200-212.

See Truong, T. K.

Edwards, C. D.

42-91 Short Baseline Phase Delay Interferometry, pp. 46-56.

Foster, C.

42-91 Analysis of the ICE Combiner for Multiple Antenna Arraying, pp. 269-277.

C. Foster and M. Marina

Fowler, L.

42-91 Traveling-Wave Maser Closed-Cycle Refrigerator Data Acquisition and Display System, pp. 304-311.

L. Fowler and M. Britcliffe

Franco, M. M.

42-89 DSS 13 Frequency Stability Tests, pp. 1-20.

See Otoshi, T. Y.

Gilchrist, C. E.

42-89 Spacecraft Mass Trade-Offs Versus Radio-Frequency Power and Antenna Size at 8 GHz and 32 GHz, pp. 21-33.

Glenn, P.

42-92 The Design and Implementation of the Technical Facilities Controller (TFC) for the Goldstone Deep Space Communications Complex, pp. 175-183.

See Killian, D. A.

Gorman, T.

42-92 The Design and Implementation of the Technical Facilities Controller (TFC) for the Goldstone Deep Space Communications Complex, pp. 175-183.

See Killian, D. A.

Gosline, R. M.

42-92 SETI Data Controllers, pp. 184-190.

Greenhall, C. A.

42-90 A Method for Using a Time Interval Counter to Measure Frequency Stability, pp. 149-156.

Guiar, C. N.

42-92 Using Ridge Regression in Systematic Pointing Error Corrections, pp. 20-26.

Gulkis, S.

42-92 Simulated Performance of an Order Statistic Threshold Strategy for Detection of Narrowband Signals, pp. 191-197.

See Satorius, E.

Hartop, R.

42-89 Microwave Component Time Delays for the 70-Meter Antennas, pp. 73-76.

42-90 Compound-Taper Feedhorn for the DSN 70-Meter Antennas, pp. 81-88.

See Manshadi, F.

Hickey, N.

42-91 A Procedural Method for the Efficient Implementation of Full-Custom VLSI Designs, pp. 235-246.

See Belk, P.

Hill, R. E.

42-91 A New State Space Model for the NASA/JPL 70-Meter Antenna Servo Controls, pp. 247-264.

- 42-91 A Modern Control Theory Based Algorithm for Control of the NASA/JPL 70-Meter Antenna Axis Servos, pp. 285-294.
- Hoppe, D.**
- 42-90 Calculated 70-Meter Antenna Performance for Offset L-Band and C-Band Feeds, pp. 12-20.
D. Hoppe and P. Stanton
- Howard, S. D.**
- 42-90 State Transition Storyboards: A Tool for Designing the Goldstone Solar System Radar Data Acquisition System User Interface Software, pp. 94-102.
- Hsu, I.-S.**
- 42-89 A VLSI Single Chip (255, 223) Reed-Solomon Encoder With Interleaver, pp. 41-47.
I.-S. Hsu, L. J. Deutsch, T. K. Truong, and I. S. Reed
- 42-90 A Comparison of VLSI Architecture of Finite Field Multipliers Using Dual, Normal, or Standard Basis, pp. 63-75.
I.-S. Hsu, T. K. Truong, H. M. Shao, L. J. Deutsch, and I.S. Reed
- 42-91 The Design Plan of a VLSI Single Chip (255, 223) Reed-Solomon Decoder, pp. 186-199.
I.-S. Hsu, H. M. Shao, and L. J. Deutsch
- 42-91 A Simplified Procedure for Correcting Both Errors and Erasures of a Reed Solomon Code Using the Euclidean Algorithm, pp. 200-212.
See Truong, T. K.
- 42-92 A Comparison of VLSI Architectures for Time and Transform Domain Decoding of Reed-Solomon Codes, pp. 63-81.
I.-S. Hsu, T. K. Truong, L. J. Deutsch, E.H. Satorius, and I. S. Reed
- Hurd, W. J.**
- 42-89 DSN Advanced Receiver: Breadboard Description and Test Results, pp. 48-66.
See Brown, D. H.
- 42-91 Optimized Tracking of RF Carriers With Phase Noise, Including Pioneer 10 Results, pp. 141-157.
See Vlnrotter, V. A.
- 42-91 Detection of Signals by the Digital Integrate-and-Dump Filter With Offset Sampling, pp. 158-173.
See Sadr, R.
- Imbriale, W. A.**
- 42-90 DSS 43 Antenna Gain Analysis for Voyager Uranus Encounter: 8.45 GHz Radio Science Data Correction, pp. 127-135.
See Slobin, S. D.
- Jackson, E. B.**
- 42-91 Power Density Measurements in the Near Field of the DSS 13 26-Meter Antenna, pp. 325-327.
E. B. Jackson and M. J. Klein
- Jee, J. R.**
- 42-91 Robust Statistical Methods for Automated Outlier Detection, pp. 23-28.
- Johns, C. E.**
- 42-89 Block III X-Band Receiver-Exciter, pp. 83-93.
- 42-91 X-Band Uplink Ground Systems Development: Part II, pp. 265-268.
- Johnson, D. L.**
- 42-91 A Cold Ejector for Closed-Cycle Helium Refrigerators, pp. 102-111.
D. L. Johnson and D. L. Daggett
- Jones, J. A.**
- 42-90 Sorption Compressor/Mechanical Expander Hybrid Refrigeration, pp. 21-26.
J. A. Jones and M. Britcliffe
- Kerr, E. L.**
- 42-91 A Near-Earth Optical Communications Terminal With a Corevolving Planetary Sun Shield, pp. 133-140.
- Killian, D. A.**
- 42-92 The Design and Implementation of the Technical Facilities Controller (TFC) for the Goldstone Deep Space Communications Complex, pp. 175-183.
D. A. Killian, F. J. Menninger, T. Gorman, and P. Glenn

Kirk, A.

- 42-89 Results of Using the Global Positioning System to Maintain the Time and Frequency Synchronization in the Deep Space Network, pp. 67-72.

See Clements, P. A.

Klein, M. J.

- 42-91 Power Density Measurements in the Near Field of the DSS 13 26-Meter Antenna, pp. 325-327.

See Jackson, E. B.

Lahmeyer, C. R.

- 42-90 Survey of Current Correlators and Applications, pp. 54-62.

C. R. Lahmeyer and S. S. Brokl

Lichten, S. M.

- 42-91 A Demonstration of High Precision GPS Orbit Determination for Geodetic Applications, pp. 1-22.

S. M. Lichten and J. S. Border

Linfield, R.

- 42-92 Source and Event Selection for Radio-Planetary Frame-Tie Measurements Using the Phobos Landers, pp. 1-12.

R. Linfield and J. Ulvestad

Loveland, R.

- 42-90 Power Supply Conditioning Circuit, pp. 1-7.

See Primas, L. E.

Lutes, G.

- 42-90 A High-Performance Single-Mode Fiber-Optic Isolator Assembly, pp. 8-11.

Maleki, L.

- 42-91 Atomic Frequency Standards for Ultra-High-Frequency Stability, pp. 67-72.

L. Maleki, J. D. Prestage, and G. J. Dick

- 42-92 The JPL Trapped Mercury Ion Frequency Standard, pp. 13-19.

See Prestage, J. D.

Manshadi, F.

- 42-90 Compound-Taper Feedhorn for the DSN 70-Meter Antennas, pp. 81-88.

F. Manshadi and R. Hartop

Marina, M.

- 42-91 Analysis of the ICE Combiner for Multiple Antenna Arraying, pp. 269-277.

See Foster, C.

McEliece, R. J.

- 42-90 Construction for Finite-State Codes, pp. 42-49.

See Pollara, F.

- 42-92 Further Results on Finite-State Codes, pp. 56-62.

See Pollara, F.

Menninger, F. J.

- 42-92 The Design and Implementation of the Technical Facilities Controller (TFC) for the Goldstone Deep Space Communications Complex, pp. 175-183.

See Killian, D. A.

Metscher, B. D.

- 42-89 An Evaluation of the Communication System for the TAU Mission Study, pp. 101-107.

Neff, D.

- 42-89 Use of a 2.3-GHz Traveling-Wave Maser on the Usuda 64-Meter Antenna, pp. 34-40.

Nelson, M. D.

- 42-92 Optical Links in the Angle-Data Assembly of the 70-Meter Antennas, pp. 154-165.

M. D. Nelson, J. R. Schroeder, and E. F. Tubbs

Nickerson, J. A.

- 42-90 A New Linear Quadratic Optimal Controller for the 34-Meter High Efficiency Antenna Position Loop, pp. 136-148.

- 42-92 New Mode Switching Algorithm for the JPL 70-Meter Antenna Servo Controller, pp. 147-153.

Norris, D.

- 42-91 Low-Noise Cryogenic Transmission Line, pp. 89-93.

Otoshi, T. Y.

- 42-89 DSS 13 Frequency Stability Tests, pp. 1-20.
T. Y. Otoshi and M. M. Franco

Pitt III, G. H.

- 42-90 Image Statistics Decoding for Convolutional Codes, pp. 32-41.
G. H. Pitt III, L. Swanson, and J. H. Yuen
- 42-92 Frame Synchronization Performance and Analysis, pp. 110-116.
See Aguilera, C. S. R.

Pollara, F.

- 42-90 Construction for Finite-State Codes, pp. 42-49.
F. Pollara, R. J. McEliece, and K. Abdel-Ghaffar
- 42-90 Soft-Decision Decoding of Some Block Codes, pp. 76-80.
- 42-92 Further Results on Finite-State Codes, pp. 56-62.
F. Pollara, K.-M. Cheung, and R. McEliece

Prestage, J. D.

- 42-91 Atomic Frequency Standards for Ultra-High-Frequency Stability, pp. 67-72.
See Maleki, L.
- 42-92 The JPL Trapped Mercury Ion Frequency Standard, pp. 13-19.
J. D. Prestage, G. J. Dick, and L. Maleki

Primas, L. E.

- 42-90 Power Supply Conditioning Circuit, pp. 1-7.
L. E. Primas and R. Loveland

Quinn, R. B.

- 42-89 Dual-Polarization 8.45 GHz Traveling-Wave Maser, pp. 94-100.
- 42-92 A 2.3-GHz Maser at Usuda, Japan, for TDRSS-Orbiting VLBI Experiment, pp. 40-42.

Reed, I. S.

- 42-89 A VLSI Single Chip (255, 223) Reed-Solomon Encoder With Interleaver, pp. 41-47.
See Hsu, I.-S.
- 42-90 A Comparison of VLSI Architecture of Finite Field Multipliers Using Dual, Normal, or Standard Basis, pp. 63-75.
See Hsu, I.-S.
- 42-91 A Simplified Procedure for Correcting Both Errors and Erasures of a Reed-Solomon Code Using the Euclidean Algorithm, pp. 200-212.
See Truong, T. K.

- 42-91 On the VLSI Design of a Pipeline Reed-Solomon Decoder Using Systolic Arrays, pp. 224-234.

See Shao, H. M.

- 42-92 A Comparison of VLSI Architectures for Time and Transform Domain Decoding of Reed-Solomon Codes, pp. 63-81.

See Hsu, I.-S.

Resch, G. M.

- 42-92 X-Band System Performance of the Very Large Array, pp. 123-137.
See Ulvestad, J. S.

Robinson, D. L.

- 42-91 Frequency Doubling Conversion Efficiencies for Deep Space Optical Communications, pp. 112-123.
D. L. Robinson and R. L. Shelton

Sadr, R.

- 42-91 Detection of Signals by the Digital Integrate-and-Dump Filter With Offset Sampling, pp. 158-173.
R. Sadr and W. J. Hurd
- 42-91 Detection of Signals by Weighted Integrate-and-Dump Filter, pp. 174-185.

Sakamoto, L. L.

- 42-92 Spacecraft Telecommunications System Mass Estimates, pp. 82-88.
See Yuen, J. H.

Satorius, E. H.

- 42-92 A Comparison of VLSI Architectures for Time and Transform Domain Decoding of Reed-Solomon Codes, pp. 63-81.

See Hsu, I.-S.

- 42-92 Simulated Performance of an Order Statistic Threshold Strategy for Detection of Narrowband Signals, pp. 191-197.

E. Satorius, R. Brady, W. Deich, S. Gulakis and E. Olsen

Schlutsmeyer, A.

- 42-89 An Interactive Wire-Wrap Board Layout Program, pp. 101-107.

Schredder, J. M.

- 42-92 Seventy-Meter Antenna Performance Predictions: GTD Analysis Compared With Traditional Ray-Tracing Methods, pp. 166-174.

Schroeder, J. R.

- 42-92 Optical Links in the Angle-Data Assembly of the 70-Meter Antennas, pp. 154-165.

See Nelson, M. D.

Shahshahani, M.

- 42-90 Systolic Arrays and Stack Decoding, pp. 50-53.

Shao, H. M.

- 42-90 A Comparison of VLSI Architecture of Finite Field Multipliers Using Dual, Normal, or Standard Basis, pp. 63-75.

See Hsu, I.-S.

- 42-91 The Design Plan of a VLSI Single Chip (255, 223) Reed-Solomon Decoder, pp. 186-199.

See Hsu, I.-S.

- 42-91 On the VLSI Design of a Pipeline Reed-Solomon Decoder Using Systolic Arrays, pp. 224-234.

H. M. Shao, L. J. Deutsch, and I. S. Reed

Shelton, R. L.

- 42-91 Frequency Doubling Conversion Efficiencies for Deep Space Optical Communications, pp. 112-123.

See Robinson, D. L.

Sigman, E. H.

- 42-92 Phase Calibration Generator, pp. 89-104.

Slobin, S. D.

- 42-90 DSS 63 64-Meter Antenna S- and X-Band Efficiency and System Noise Temperature Calibrations, July 1986, pp. 103-115.

- 42-90 DSS 43 64-Meter Antenna S- and X-Band Efficiency and System Noise Temperature Calibrations, January 1987, pp. 116-126.

- 42-90 DSS 43 Antenna Gain Analysis for Voyager Uranus Encounter: 8.45 GHz Radio Science Data Correction, pp. 127-135.

S. D. Slobin and W. A. Imbriale

- 42-92 DSS 14 64-Meter Antenna S- and X-Band Efficiency and System Noise Temperature Calibrations, September 1987, pp. 138-146.

Smith, J. R.

- 42-92 A Simplified, General-Purpose Deep-Space Ranging Correlator Design, pp. 105-109.

See Tausworthe, R. C.

Stanton, P. H.

- 42-90 Calculated 70-Meter Antenna Performance for Offset L-Band and C-Band Feeds, pp. 12-20.

See Hoppe, D.

- 42-90 X-Band Phase Calibration Generator Coupler, pp. 89-93.

Stelzried, C. T.

- 42-91 Non-Linearity in Measurement Systems: Evaluation Method and Application to Microwave Radiometers, pp. 57-66.

Swanson, L.

- 42-90 Image Statistics Decoding for Convolutional Codes, pp. 32-41.

See Pitt III, G. H.

- 42-92 Frame Synchronization Performance and Analysis, pp. 110-116.

See Aguilera, C. S. R.

Tanida, L.

- 42-91 A 2.3-GHz Cryogenically Cooled HEMT Amplifier for DSS 13, pp. 94-101.

Tausworthe, R. C.

- 42-91 Tau Ranging Revisited, pp. 318-324.
- 42-92 A Simplified, General-Purpose Deep-Space Ranging Correlator Design, pp. 105-109.
R. C. Tausworthe and J. R. Smith

Truong, T. K.

- 42-89 A VLSI Single Chip (255, 223) Reed-Solomon Encoder With Interleaver, pp. 41-47.
See Hsu, I.-S.
- 42-90 A Comparison of VLSI Architecture of Finite Field Multipliers Using Dual, Normal, or Standard Basis, pp. 63-75.
See Hsu, I.-S.
- 42-91 A Simplified Procedure for Correcting Both Errors and Erasures of a Reed-Solomon Code Using the Euclidean Algorithm, pp. 200-212.
T. K. Truong, I. S. Hsu, W. L. Eastman, and I. S. Reed
- 42-92 A Comparison of VLSI Architectures for Time and Transform Domain Decoding of Reed-Solomon Codes, pp. 63-81.
See Hsu, I.-S.

Tubbs, E. F.

- 42-92 Optical Links in the Angle-Data Assembly of the 70-Meter Antennas, pp. 154-165.
See Nelson, M. D.

Tucker, T. K.

- 42-91 Fast Autotuning of a Hydrogen Maser by Cavity Q Modulation, pp. 295-303.
See Dick, G. J.

Ulvestad, J. S.

- 42-92 Source and Event Selection for Radio-Planetary Frame-Tie Measurements Using the Phobos Landers, pp. 1-12.
See Linfield, R.

- 42-92 X-Band System Performance of the Very Large Array, pp. 123-137.

J. S. Ulvestad, G. M. Resch, and W. D. Brundage

Unglaub, R.

- 42-89 Results of Using the Global Positioning System to Maintain the Time and Frequency Synchronization in the Deep Space Network, pp. 67-72.
See Clements, P.A.

Vilnrotter, V. A.

- 42-91 Optimized Tracking of RF Carriers With Phase Noise, Including Pioneer 10 Results, pp. 141-157.
V. A. Vilnrotter, W. J. Hurd, and D. H. Brown

Wu, S. C.

- 42-91 Precise Near-Earth Navigation With GPS: A Survey of Techniques, pp. 29-45.
See Yunck, T. P.

Wu, J.

- 42-91 Precise Near-Earth Navigation With GPS: A Survey of Techniques, pp. 29-45.
See Yunck, T. P.

Yuen, J. H.

- 42-90 Image Statistics Decoding for Convolutional Codes, pp. 32-41.
See Pitt III, G. H.
- 42-92 Spacecraft Telecommunications System Mass Estimates, pp. 82-88.
J. H. Yuen and L. L. Sakamoto

Yunck, T. P.

- 42-91 Precise Near-Earth Navigation With GPS: A Survey of Techniques, pp. 29-45.
T. P. Yunck, S. C. Wu, and J. Wu